STIC Search Report
Biotech-Chem Library

STIC Database Tracking Number: 121508

TO: Jennifer Graser **Location: rem/3b09/3c18** **Art Unit: 1645** **Monday, May 10, 2004** **Case Serial Number: 10/625221**	**From: Edward Hart** **Location: Biotech-Chem Library** **REM-1A55** **Phone: 571-272-2512** **edward.hart@uspto.gov**

Search Notes

Examiner Graser,

Here are the results of the search you requested.

Please feel free to contact me if you have any questions.

Edward Hart

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: May 7, 2004, 12:05:23 ; Search time 21 Seconds
(without alignments)
1012.301 Million cell updates/sec

Title: US-10-625-221-14
Perfect score: 1178
Sequence: 1 QQDPDPSQLHRSSLVKNLQN.....KDNETLDSNTSQIEVYLTTK 221

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 283366 seqs, 96191526 residues

Total number of hits satisfying chosen parameters: 283366

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : PIR_78:*
1: pir1:*
2: pir2:*
3: pir3:*
4: pir4:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	1169	99.2	251	1	S29659	exotoxin type A pr
2	1133	96.2	236	2	S18783	exotoxin type A pr
3	1128	95.8	236	2	S18786	exotoxin type A pr
4	1025	87.0	236	2	S18789	exotoxin A precurs
5	978.5	83.1	250	1	A26152	streptococcal pyro
6	572	48.6	266	1	ENSAB6	enterotoxin B prec
7	546	46.3	266	2	S11885	enterotoxin C3 - S
8	543	46.1	266	2	A60114	enterotoxin C-2 pr
9	542	46.0	266	1	ENSAC1	enterotoxin C-1 pr
10	468	39.7	258	2	G89968	extracellular ente
11	337	28.6	136	2	A89969	enterotoxin YENT2
12	333.5	28.3	258	2	A33953	enterotoxin D prec
13	323.5	27.5	260	2	E89969	enterotoxin SeO [i
14	315.5	26.8	260	2	C89984	enterotoxin P [imp
15	297.5	25.3	258	2	H89968	enterotoxin SeN [i
16	293.5	24.9	257	2	A28664	enterotoxin A prec
17	291.5	24.7	257	2	A28179	enterotoxin E prec
18	284	24.1	240	2	G89991	extracellular ente
19	265.5	22.5	239	2	D89969	enterotoxin SEM [i
20	229.5	19.5	242	2	C89969	extracellular ente
21	205	17.4	235	2	A30509	exotoxin C precurs
22	173	14.7	133	2	B89969	enterotoxin Yent1
23	127	10.8	231	2	D89807	exotoxin 11 [impor
24	123.5	10.5	157	2	A89942	hypothetical prote
25	123	10.4	227	2	C89808	exotoxin 15 [impor
26	114	9.7	292	2	B89807	exotoxin 9 [import
27	112.5	9.6	234	2	C89807	exotoxin 10 [impor
28	109.5	9.3	1367	2	T18466	hypothetical prote
29	105.5	9.0	232	2	F89807	exotoxin 13 [impor
30	105.5	9.0	596	2	E96935	FUSION, methionine
31	105.5	9.0	825	2	H82885	hypothetical prote
32	102	8.7	1856	2	C95008	immunoglobulin A1
33	101	8.6	234	1	XCSAS1	toxic shock syndro
34	101	8.6	770	2	B48910	desmocollin 1b pre
35	101	8.6	824	2	A48910	desmocollin 1a pre
36	101	8.6	840	2	I37281	Dsc1a precursor -
37	101	8.6	894	2	I37282	Dsc1b precursor -
38	100	8.5	4981	2	T18489	hypothetical prote
39	98	8.3	234	2	B89992	toxic shock syndro
40	98	8.3	493	2	G90604	hypothetical prote
41	97.5	8.3	463	2	D95019	hypothetical prote
42	96	8.1	241	2	B89888	hypothetical prote
43	96	8.1	396	2	S09627	prrC protein - Esc
44	96	8.1	908	2	C70168	DNA polymerase I (
45	96	8.1	1306	2	T28313	ORF MSV152 probabl

ALIGNMENTS

RESULT 1
S29659
exotoxin type A precursor (allele 1) - Streptococcus pyogenes phage T12
N;Alternate names: erythrogenic toxin; scarlet fever toxin
C;Species: Streptococcus pyogenes phage T12
C;Date: 10-Sep-1999 #sequence_revision 10-Sep-1999 #text_change 10-Sep-1999
C;Accession: S29659; S18782; S18784; S18785; S18791; S18796; S18797; S18800
R;Weeks, C.R.; Ferretti, J.J.
Infect. Immun. 52, 144-150, 1986
A;Title: Nucleotide sequence of the type A streptococcal exotoxin (erythrogenic toxin) g
A;Reference number: S29659; MUID:86166804; PMID:3514452
A;Accession: S29659
A;Molecule type: DNA
A;Residues: 1-251 <WEE>
A;Cross-references: GB:U40453; EMBL:M19350; NID:g1877426; PIDN:AAC48868.1; PID:g1877430
R;Nelson, K.; Schlievert, P.M.; Selander, R.K.; Musser, J.M.
J. Exp. Med. 174, 1271-1274, 1991
A;Title: Characterization and clonal distribution of four alleles of the speA gene encod
A;Reference number: S18782; MUID:92044323; PMID:1940804
A;Accession: S18782
A;Status: nucleic acid sequence not shown; translation not shown
A;Molecule type: DNA
A;Residues: 9-244 <NEL>
A;Cross-references: EMBL:X61560; NID:g47287; PIDN:CAA43758.1; PID:g47288
A;Experimental source: Streptococcus pyogenes strain MGAS156 isolate Nebraska unassigned
A;Note: the nucleotide sequence was submitted to the EMBL Data Library, September 1991
A;Accession: S18784
A;Status: nucleic acid sequence not shown; translation not shown
A;Molecule type: DNA
A;Residues: 9-244 <NEA>
A;Cross-references: EMBL:X61556; NID:g47291; PIDN:CAA43754.1; PID:g47292
A;Experimental source: Streptococcus pyogenes strain MGAS165 isolate Minnesota unassigne
A;Note: the nucleotide sequence was submitted to the EMBL Data Library, September 1991
A;Accession: S18785
A;Status: nucleic acid sequence not shown; translation not shown
A;Molecule type: DNA
A;Residues: 9-244 <NEZ>
A;Cross-references: EMBL:X61559; NID:g47293; PIDN:CAA43757.1; PID:g47294
A;Experimental source: Streptococcus pyogenes strain MGAS167 isolate Texas unassigned ph
A;Note: the nucleotide sequence was submitted to the EMBL Data Library, September 1991
A;Accession: S18791
A;Status: nucleic acid sequence not shown; translation not shown
A;Molecule type: DNA
A;Residues: 9-244 <NEY>
A;Cross-references: EMBL:X61555; NID:g47309; PIDN:CAA43753.1; PID:g47310
A;Experimental source: Streptococcus pyogenes strain MGAS327 isolate Arizona unassigned
A;Note: the nucleotide sequence was submitted to the EMBL Data Library, September 1991
A;Accession: S18796
A;Status: nucleic acid sequence not shown; translation not shown
A;Molecule type: DNA
A;Residues: 9-244 <NEO>
A;Cross-references: EMBL:X61557; NID:g47319; PIDN:CAA43755.1; PID:g47320

A;Experimental source: Streptococcus pyogenes strain MGAS493 isolate France unassigned p
A;Note: the nucleotide sequence was submitted to the EMBL Data Library, September 1991
A;Accession: S18797
A;Status: nucleic acid sequence not shown; translation not shown
A;Molecule type: DNA
A;Residues: 9-244 <NEH>
A;Cross-references: EMBL:X61558; NID:g47321; PIDN:CAA43756.1; PID:g47322
A;Experimental source: Streptococcus pyogenes strain MGAS494 isolate France unassigned p
A;Note: the nucleotide sequence was submitted to the EMBL Data Library, September 1991
A;Accession: S18800
A;Status: nucleic acid sequence not shown; translation not shown
A;Molecule type: DNA
A;Residues: 9-228 <NES>
A;Cross-references: EMBL:X61554; NID:g47327; PIDN:CAA43752.1; PID:g47328
A;Experimental source: Streptococcus pyogenes strain MGAS500 isolate New Zealand unassig
A;Note: the nucleotide sequence was submitted to the EMBL Data Library, September 1991
C;Genetics:
A;Gene: speA; speA1
C;Superfamily: enterotoxin B
C;Keywords: exotoxin
F;1-30/Domain: signal sequence #status predicted <SIG>
F;31-251/Product: exotoxin type A #status predicted <MAT>

Query Match 99.2%; Score 1169; DB 1; Length 251;
Best Local Similarity 99.5%; Pred. No. 4e-83;
Matches 220; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

```
Qy          1 QQDPDPSQLHRSSLVKNLQNIYFLYEGDPVTHENVKSVDQLLSHHLIYNVSGPNYDKLKT 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         31 QQDPDPSQLHRSSLVKNLQNIYFLYEGDPVTHENVKSVDQLLSHDLIYNVSGPNYDKLKT 90

Qy         61 ELKNQEMATLFKDKNVDIYGVEYYHLCYLCENAERSACIYGGVTNHEGNHLEIPKKIVVK 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         91 ELKNQEMATLFKDKNVDIYGVEYYHLCYLCENAERSACIYGGVTNHEGNHLEIPKKIVVK 150

Qy        121 VSIDGIQSLSFDIETNKKMVTAQELDYKVRKYLTDNKQLYTNGPSKYETGYIKFIPKNKE 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        151 VSIDGIQSLSFDIETNKKMVTAQELDYKVRKYLTDNKQLYTNGPSKYETGYIKFIPKNKE 210

Qy        181 SFWFDFFPEPEFTQSKYLMIYKDNETLDSNTSQIEVYLTTK 221
              |||||||||||||||||||||||||||||||||||||||||||
Db        211 SFWFDFFPEPEFTQSKYLMIYKDNETLDSNTSQIEVYLTTK 251
```

RESULT 2
S18783
exotoxin type A precursor (allele 3) - Streptococcus pyogenes phage (strain MGAS158 isol
N;Alternate names: scarlet fever toxin
C;Species: Streptococcus pyogenes phage
A;Variety: strain MGAS158 isolate Nebraska; strain MGAS485 isolate Yugoslavia; strain MG
C;Date: 29-Jan-1993 #sequence_revision 29-Jan-1993 #text_change 16-Jul-1999
C;Accession: S18783; S18793; S18794; S18801; S18798
R;Nelson, K.; Schlievert, P.M.; Selander, R.K.; Musser, J.M.
J. Exp. Med. 174, 1271-1274, 1991
A;Title: Characterization and clonal distribution of four alleles of the speA gene encod
A;Reference number: S18782; MUID:92044323; PMID:1940804
A;Accession: S18783
A;Status: nucleic acid sequence not shown; translation not shown
A;Molecule type: DNA
A;Residues: 1-236 <NEL>
A;Cross-references: EMBL:X61568; NID:g47289; PIDN:CAA43766.1; PID:g47290
A;Experimental source: strain MGAS158 isolate Nebraska unassigned phage
A;Note: the nucleotide sequence was submitted to the EMBL Data Library, September 1991
A;Accession: S18793
A;Status: nucleic acid sequence not shown; translation not shown
A;Molecule type: DNA
A;Residues: 1-236 <NEA>
A;Cross-references: EMBL:X61569; NID:g47313; PIDN:CAA43767.1; PID:g47314
A;Experimental source: strain MGAS485 isolate Yugoslavia unassigned phage
A;Note: the nucleotide sequence was submitted to the EMBL Data Library, September 1991
A;Accession: S18794
A;Status: nucleic acid sequence not shown; translation not shown
A;Molecule type: DNA
A;Residues: 1-236 <NEZ>
A;Cross-references: EMBL:X61570; NID:g47315; PIDN:CAA43768.1; PID:g47316
A;Experimental source: strain MGAS491 isolate United Kingdom unassigned phage
A;Note: the nucleotide sequence was submitted to the EMBL Data Library, September 1991
A;Accession: S18801
A;Status: nucleic acid sequence not shown; translation not shown
A;Molecule type: DNA
A;Residues: 1-236 <NEY>
A;Cross-references: EMBL:X61572; NID:g47333; PIDN:CAA43770.1; PID:g47334
A;Experimental source: strain MGAS624 isolate Germany unassigned phage
A;Note: the nucleotide sequence was submitted to the EMBL Data Library, September 1991
A;Accession: S18798
A;Status: nucleic acid sequence not shown; translation not shown
A;Molecule type: DNA
A;Residues: 1-236 <NEO>
A;Cross-references: EMBL:X61571; NID:g47323; PIDN:CAA43769.1; PID:g47324
A;Experimental source: strain MGAS495 isolate Germany unassigned phage
A;Note: the nucleotide sequence was submitted to the EMBL Data Library, September 1991
C;Genetics:
A;Gene: speA3
C;Superfamily: enterotoxin B
C;Keywords: exotoxin
F;1-22/Domain: signal sequence (fragment) #status predicted <SIG>
F;23-236/Product: exotoxin type A (fragment) #status predicted <MAT>

Query Match 96.2%; Score 1133; DB 2; Length 236;
Best Local Similarity 99.1%; Pred. No. 2.2e-80;
Matches 212; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

```
Qy          1 QQDPDPSQLHRSSLVKNLQNIYFLYEGDPVTHENVKSVDQLLSHHLIYNVSGPNYDKLKT 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         23 QQDPDPSQLHRSSLVKNLQNIYFLYEGDPVTHENVKSVDQLLSHDLIYNVSGPNYDKLKT 82

Qy         61 ELKNQEMATLFKDKNVDIYGVEYYHLCYLCENAERSACIYGGVTNHEGNHLEIPKKIVVK 120
              ||||||||||||||:||||||||||||||||||||||||||||||||||||||||||||||
Db         83 ELKNQEMATLFKDKNIDIYGVEYYHLCYLCENAERSACIYGGVTNHEGNHLEIPKKIVVK 142

Qy        121 VSIDGIQSLSFDIETNKKMVTAQELDYKVRKYLTDNKQLYTNGPSKYETGYIKFIPKNKE 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        143 VSIDGIQSLSFDIETNKKMVTAQELDYKVRKYLTDNKQLYTNGPSKYETGYIKFIPKNKE 202

Qy        181 SFWFDFFPEPEFTQSKYLMIYKDNETLDSNTSQI 214
              |||||||||||||||||||||||||||||||||||
Db        203 SFWFDFFPEPEFTQSKYLMIYKDNETLDSNTSQI 236
```

RESULT 3
S18786
exotoxin type A precursor (allele 2) - Streptococcus pyogenes phage (strain MGAS250 isola
N;Alternate names: scarlet fever toxin
C;Species: Streptococcus pyogenes phage
A;Variety: strain MGAS250 isolate California; strain MGAS251 isolate California; strain M
isolate United Kingdom; strain MGAS496 isolate Germany
C;Date: 29-Jan-1993 #sequence_revision 29-Jan-1993 #text_change 16-Jul-1999
C;Accession: S18786; S18787; S18788; S18790; S18792; S18795; S18799
R;Nelson, K.; Schlievert, P.M.; Selander, R.K.; Musser, J.M.
J. Exp. Med. 174, 1271-1274, 1991
A;Title: Characterization and clonal distribution of four alleles of the speA gene encod
A;Reference number: S18782; MUID:92044323; PMID:1940804
A;Accession: S18786
A;Status: nucleic acid sequence not shown; translation not shown
A;Molecule type: DNA
A;Residues: 1-236 <NEL>
A;Cross-references: EMBL:X61561; NID:g47297; PIDN:CAA43759.1; PID:g47298
A;Experimental source: strain MGAS250 isolate California unassigned phage
A;Note: the nucleotide sequence was submitted to the EMBL Data Library, September 1991
A;Accession: S18787
A;Status: nucleic acid sequence not shown; translation not shown
A;Molecule type: DNA
A;Residues: 1-236 <NEA>
A;Cross-references: EMBL:X61562; NID:g47299; PIDN:CAA43760.1; PID:g47300
A;Experimental source: strain MGAS251 isolate California unassignd phage
A;Note: the nucleotide sequence was submitted to the EMBL Data Library, September 1991

A;Accession: S18788
A;Status: nucleic acid sequence not shown; translation not shown
A;Molecule type: DNA
A;Residues: 1-236 <NEZ>
A;Cross-references: EMBL:X61563; NID:g47301; PIDN:CAA43761.1; PID:g47302
A;Experimental source: strain MGAS256 isolate California unassigned phage
A;Note: the nucleotide sequence was submitted to the EMBL Data Library, September 1991
A;Accession: S18790
A;Status: nucleic acid sequence not shown; translation not shown
A;Molecule type: DNA
A;Residues: 1-236 <NEY>
A;Cross-references: EMBL:X61564; NID:g47305; PIDN:CAA43762.1; PID:g47306
A;Experimental source: strain MGAS285 isolate Colorado unassigned phage
A;Note: the nucleotide sequence was submitted to the EMBL Data Library, September 1991
A;Accession: S18792
A;Status: nucleic acid sequence not shown; translation not shown
A;Molecule type: DNA
A;Residues: 1-236 <NEO>
A;Cross-references: EMBL:X61565; NID:g47311; PIDN:CAA43763.1; PID:g47312
A;Experimental source: strain MGAS480 isolate Yugoslavia unassigned phage
A;Note: the nucleotide sequence was submitted to the EMBL Data Library, September 1991
A;Accession: S18795
A;Status: nucleic acid sequence not shown; translation not shown
A;Molecule type: DNA
A;Residues: 1-236 <NEH>
A;Cross-references: EMBL:X61566; NID:g47317; PIDN:CAA43764.1; PID:g47318
A;Experimental source: strain MGAS492 isolate United Kingdom unassigned phage
A;Note: the nucleotide sequence was submitted to the EMBL Data Library, September 1991
A;Accession: S18799
A;Status: nucleic acid sequence not shown; translation not shown
A;Molecule type: DNA
A;Residues: 1-236 <NES>
A;Cross-references: EMBL:X61567; NID:g47325; PIDN:CAA43765.1; PID:g47326
A;Experimental source: strain MGAS496 isolate Germany unassigned phage
A;Note: the nucleotide sequence was submitted to the EMBL Data Library, September 1991
C;Genetics:
A;Gene: speA2
C;Superfamily: enterotoxin B
C;Keywords: exotoxin
F;1-22/Domain: signal sequence (fragment) #status predicted <SIG>
F;23-236/Product: exotoxin type A (fragment) #status predicted <MAT>

Query Match 95.8%; Score 1128; DB 2; Length 236;
Best Local Similarity 99.1%; Pred. No. 5.4e-80;
Matches 212; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

```
Qy          1 QQDPDPSQLHRSSLVKNLQNIYFLYEGDPVTHENVKSVDQLLSHHLIYNVSGPNYDKLKT 60
              ||||||||||||||||||||||||||||||||||||||||||| ||||||||||||||||
Db         23 QQDPDPSQLHRSSLVKNLQNIYFLYEGDPVTHENVKSVDQLLSHDLIYNVSGPNYDKLKT 82

Qy         61 ELKNQEMATLFKDKNVDIYGVEYYHLCYLCENAERSACIYGGVTNHEGNHLEIPKKIVVK 120
              |||||||||||||||||| ||||||||||||||||||||||||||||||||||||||||
Db         83 ELKNQEMATLFKDKNVDIYSVEYYHLCYLCENAERSACIYGGVTNHEGNHLEIPKKIVVK 142

Qy        121 VSIDGIQSLSFDIETNKKMVTAQELDYKVRKYLTDNKQLYTNGPSKYETGYIKFIPKNKE 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        143 VSIDGIQSLSFDIETNKKMVTAQELDYKVRKYLTDNKQLYTNGPSKYETGYIKFIPKNKE 202

Qy        181 SFWFDFFPEPEFTQSKYLMIYKDNETLDSNTSQI 214
              |||||||||||||||||||||||||||||||||||
Db        203 SFWFDFFPEPEFTQSKYLMIYKDNETLDSNTSQI 236
```

RESULT 4
S18789
exotoxin A precursor (allele 4) - Streptococcus pyogenes (strain MGAS262 isolate Califor
N;Alternate names: scarlet fever toxin
C;Species: Streptococcus pyogenes
A;Variety: strain MGAS262 isolate California
C;Date: 29-Jan-1993 #sequence_revision 29-Jan-1993 #text_change 16-Jul-1999
C;Accession: S18789
R;Nelson, K.; Schlievert, P.M.; Selander, R.K.; Musser, J.M.
J. Exp. Med. 174, 1271-1274, 1991
A;Title: Characterization and clonal distribution of four alleles of the speA gene encod
A;Reference number: S18782; MUID:92044323; PMID:1940804
A;Accession: S18789
A;Status: nucleic acid sequence not shown; translation not shown
A;Molecule type: DNA
A;Residues: 1-236 <NEL>
A;Cross-references: EMBL:X61573; NID:g47303; PIDN:CAA43771.1; PID:g47304
A;Note: the nucleotide sequence was submitted to the EMBL Data Library, September 1991
C;Genetics:
A;Gene: speA
C;Superfamily: enterotoxin B
C;Keywords: exotoxin
F;1-22/Domain: signal sequence (fragment) #status predicted <SIG>
F;23-236/Product: exotoxin A (fragment) #status predicted <MAT>

Query Match 87.0%; Score 1025; DB 2; Length 236;
Best Local Similarity 89.7%; Pred. No. 4.9e-72;
Matches 192; Conservative 12; Mismatches 10; Indels 0; Gaps 0;

```
Qy          1 QQDPDPSQLHRSSLVKNLQNIYFLYEGDPVTHENVKSVDQLLSHHLIYNVSGPNYDKLKT 60
              ||||:||||||||||||||||||||||| |||||||||||| ||||||| |||||||
Db         23 QQDPNPSQLHRSSLVKNLQNIYFLYEGDPVVHENVKSVDQLLSHDLIYNVSGLNYDKLKT 82

Qy         61 ELKNQEMATLFKDKNVDIYGVEYYHLCYLCENAERSACIYGGVTNHEGNHLEIPKKIVVK 120
              ||||:||:|||||:||||||||||: |||| ||:| |||||||||||||||||| |:||
Db         83 ELKNREMSTLFKNKNVDIYGVEYYYHCYLCRNAKRRACIYGGVTNHEGNHLEIPKNILVK 142

Qy        121 VSIDGIQSLSFDIETNKKMVTAQELDYKVRKYLTDNKQLYTNGPSKYETGYIKFIPKNKE 180
              ||||||||||||||:|||||||||||||||:|||||||||||||||||||||||| |:||
Db        143 VSIDGIQSLSFDIETSKKMVTAQELDYKVRKHLTDNKQLYTNGPSKYETGYIKFISKDKE 202

Qy        181 SFWFDFFPEPEFTQSKYLMIYKDNETLDSNTSQI 214
              :|||||||||| | ||||||||||||:||||
Db        203 TFWFDFFPEPEFNQVKYLMIYKDNETLDSSTSQI 236
```

RESULT 5
A26152
streptococcal pyrogenic exotoxin type A precursor - Streptococcus sp.
N;Alternate names: scarlet fever toxin; SPE type A (speA)
C;Species: Streptococcus sp.
C;Date: 10-Sep-1999 #sequence_revision 10-Sep-1999 #text_change 10-Sep-1999
C;Accession: A26152
R;Johnson, L.P.; L'Italien, J.J.; Schlievert, P.M.
Mol. Gen. Genet. 203, 354-356, 1986
A;Title: Streptococcal pyrogenic exotoxin type A (scarlet fever toxin) is related to Sta
A;Reference number: A26152; MUID:86284313; PMID:3526093
A;Accession: A26152
A;Molecule type: DNA
A;Residues: 1-250 <JOH>
C;Superfamily: enterotoxin B
C;Keywords: exotoxin

Query Match 83.1%; Score 978.5; DB 1; Length 250;
Best Local Similarity 86.7%; Pred. No. 2.1e-68;
Matches 189; Conservative 4; Mismatches 24; Indels 1; Gaps 1;

```
Qy          4 PDPSQLHRSSLVKNLQNIYFLYEGDPVTHENVKSVDQLLSHHLIYNVSGPNYDKLKTELK 63
              | |||| ||:||| :  |       |||||||||||||||| ||||||||||||||||
Db         34 PKPSQLQRSNLVKTFKIYIFFMRVTLVTHENVKSVDQLLSHDLIYNVSGPNYDKLKTELK 93

Qy         64 NQEMATLFKDKNVDIYGVEYYHLCYLCENAERSACIYGGVTNHEGNHLEIPKKIVVKVSI 123
              ||||||||||||||||||||||||||||||:|||||||||||||||||||||||||||
Db         94 NQEMATLFKDKNVDIYGVEYYHLCYLCENAERSACLYGGVTNHEGNHLEIPKKIVVKVSI 153

Qy        124 DGIQSLSFDIETNKKMVTAQELDYKVRKYLTDNKQLYTNGPSKYETGYIKFIPKNKESFW 183
              |||||||||||  |       : | ||||||||||||||||||||||||||||||||||
Db        154 DGIQSLSFDIEQIKN-GNCSRISYTVRKYLTDNKQLYTNGPSKYETGYIKFIPKNKESFW 212

Qy        184 FDFFPEPEFTQSKYLMIYKDNETLDSNTSQIEVYLTTK 221
              |||||||||||||||||||||||||||||||||||||
Db        213 FDFFPEPEFTQSKYLMIYKDNETLDSNTSQIEVYLTTK 250
```

RESULT 6
ENSAB6
enterotoxin B precursor - Staphylococcus aureus
C;Species: Staphylococcus aureus
C;Date: 24-Apr-1984 #sequence_revision 15-Oct-1996 #text_change 18-Jun-1999
C;Accession: S27360; A92065; S27240; A01815
R;Jones, C.L.; Khan, S.A.
J. Bacteriol. 166, 29-33, 1986
A;Title: Nucleotide sequence of the enterotoxin B gene from Staphylococcus aureus.
A;Reference number: S27360; MUID:86168029; PMID:3957869
A;Accession: S27360
A;Molecule type: DNA
A;Residues: 1-266 <JON>
A;Cross-references: EMBL:M11118; NID:g152999; PIDN:AAA88550.1; PID:g153000
A;Experimental source: strain S6
R;Huang, I.Y.; Bergdoll, M.S.
J. Biol. Chem. 245, 3518-3525, 1970
A;Title: The primary structure of staphylococcal enterotoxin B. III. The cyanogen bromid
A;Reference number: A92065; MUID:71007902; PMID:5470821
A;Accession: A92065
A;Molecule type: protein
A;Residues: 28-55,'NND',59-68,'NE',71,'FDLIYL',78-117,119-127,'N',129,'D',131-132,'ENT',
A;Experimental source: strain S-6
R;Huang, I.Y.; Bergdoll, M.S.
J. Biol. Chem. 245, 3511-3517, 1970
A;Title: The primary structure of staphylococcal enterotoxin B. II. Isolation, compositi
A;Reference number: A92064; MUID:71007901; PMID:5470820
A;Contents: annotation; chymotryptic peptides
R;Huang, I.Y.; Bergdoll, M.S.
J. Biol. Chem. 245, 3493-3510, 1970
A;Title: The primary structure of staphylococcal enterotoxin B. I. Isolation, compositio
A;Reference number: A92063; MUID:71007900; PMID:5470819
A;Contents: annotation; tryptic peptides
R;Schantz, E.J.; Roessler, W.G.; Wagman, J.; Spero, L.; Dunnery, D.A.; Bergdoll, M.S.
Biochemistry 4, 1011-1016, 1965
A;Title: Purification of staphylococcal enterotoxin B.
A;Reference number: A90548; MUID:66035792; PMID:4953912
A;Contents: annotation; biological source of protein
R;Alakhov, V.Y.; Klinsky, E.Y.; Kolosov, M.I.; Maurer-Fogy, I.; Moskaleva, E.Y.; Sveshni
Eur. J. Biochem. 209, 823-828, 1992
A;Title: Identification of functionally active fragments of staphylococcal enterotoxin B
A;Reference number: S27240; MUID:93049338; PMID:1425690
A;Accession: S27240
A;Molecule type: protein
A;Residues: 28-42;128-148 <ALA>
C;Superfamily: enterotoxin B
C;Keywords: enterotoxin; extracellular protein; toxin
F;1-27/Domain: signal sequence #status predicted <SIG>
F;28-266/Product: enterotoxin B #status experimental <MAT>
F;120-140/Disulfide bonds: #status experimental

Query Match 48.6%; Score 572; DB 1; Length 266;
Best Local Similarity 49.2%; Pred. No. 5.6e-37;
Matches 116; Conservative 36; Mismatches 68; Indels 16; Gaps 5;

```
Qy          1 QQDPDPSQLHRSSLVKNL-QNIYFLYEGDPVTHENVKSVDQLLSHHLIYNVSGP---NYD 56
              | || | :||:||    | :|:  ||: : |:  ||||:|| |   |||::      |||
Db         30 QPDPKPDELHKSSKFTGLMENMKVLYDDNHVSAINVKSIDQFLYFDLIYSIKDTKLGNYD 89

Qy         57 KLKTELKNQEMATLFKDKNVDIYGVEYYHLCYLCENA---------ERSACIYGGVTNHE 107
               :: | ||:::|  :||| ||::|  ||: ||  :           :|  |:||||| |
Db         90 NVRVEFKNKDLADKYKDKYVDVFGANYYYQCYFSKKTNDINSHQTDKRKTCMYGGVTEHN 149

Qy        108 GNHLEIPKKIVVKVSIDGIQSLSFDIETNKKMVTAQELDYKVRKYLTDNKQLYTNGPSKY 167
              || |:  : | |:|  ||   ||||::|||| ||||||||  | ||  ||:||     | |
Db        150 GNQLDKYRSITVRVFEDGKNLLSFDVQTNKKKVTAQELDYLTRHYLVKNKKLYEFNNSPY 209

Qy        168 ETGYIKFIPKNKESFWFDFFPEP--EFTQSKYLMIYKDNETLDSNTSQIEVYLTTK 221
              |||||||| :|: |||:|  | |  :| ||||||:| ||: :||   :||||||||
Db        210 ETGYIKFI-ENENSFWYDMMPAPGDKFDQSKYLMMYNDNKMVDSKDVKIEVYLTTK 264
```

RESULT 7
S11885
enterotoxin C3 - Staphylococcus aureus
C;Species: Staphylococcus aureus
C;Date: 19-Mar-1997 #sequence_revision 19-Mar-1997 #text_change 16-Jul-1999
C;Accession: S11885
R;Hovde, C.J.; Hackett, S.P.; Bohach, G.A.
Mol. Gen. Genet. 220, 329-333, 1990
A;Title: Nucleotide sequence of the staphylococcal enterotoxin C3 gene: sequence comparis
A;Reference number: S11885; MUID:90220508; PMID:2325627
A;Accession: S11885
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-266 <HOV>
A;Cross-references: GB:X51661; NID:g46570; PIDN:CAA35972.1; PID:g46571
C;Superfamily: enterotoxin B

Query Match 46.3%; Score 546; DB 2; Length 266;
Best Local Similarity 47.2%; Pred. No. 5.7e-35;
Matches 111; Conservative 43; Mismatches 67; Indels 14; Gaps 6;

```
Qy          1 QQDPDPSQLHRSS-LVKNLQNIYFLYEGDPVTHENVKSVDQLLSHHLIYNVSG---PNYD 56
              | || |  ||:||      : |: :||:    |:   |||||: |:| ||||:|     |||
Db         30 QPDPMPDDLHKSSEFTGTMGNMKYLYDDHYVSATKVKSVDKFLAHDLIYNISDKKLKNYD 89

Qy         57 KLKTELKNQEMATLFKDKNVDIYGVEYYHLCYLC--ENAER----SACIYGGVTNHEGNH 110
              |:|||| |:::|  :||: ||:||  ||  ||     :|  :       |:|||:| |||||
Db         90 KVKTELLNEDLAKKYKDEVVDVYGSNYYVNCYFSSKDNVGKVTGGKTCMYGGITKHEGNH 149

Qy        111 LEIP--KKIVVKVSIDGIQSLSFDIETNKKMVTAQELDYKVRKYLTDNKQLYTNGPSKYE 168
              :    : ::|:|  :   ::||:::|:|| |||||||  | | :| : | ||     | ||
Db        150 FDNGNLQNVLVRVYENKRNTISFEVQTDKKSVTAQELDIKARNFLINKKNLYEFNSSPYE 209

Qy        169 TGYIKFIPKNKESFWFDFFPEP--EFTQSKYLMIYKDNETLDSNTSQIEVYLTTK 221
              |||||||  |  :||:|  | |  :| ||||||:| ||:|:|| : :|||:||||
Db        210 TGYIKFIENNGNTFWYDMMPAPGDKFDQSKYLMMYNDNKTVDSKSVKIEVHLTTK 264
```

RESULT 8
A60114
enterotoxin C-2 precursor - Staphylococcus aureus
N;Alternate names: enterotoxin C-3 precursor
C;Species: Staphylococcus aureus
C;Date: 10-Nov-1992 #sequence_revision 10-Nov-1992 #text_change 16-Jul-1999
C;Accession: A60114; B60114; A33866
R;Bohach, G.A.; Schlievert, P.M.
Infect. Immun. 57, 2249-2252, 1989
A;Title: Conservation of the biologically active portions of staphylococcal enterotoxins
A;Reference number: A60114; MUID:89277549; PMID:2543637
A;Accession: A60114
A;Status: not compared with conceptual translation
A;Molecule type: DNA
A;Residues: 1-266 <BOH>
A;Accession: B60114
A;Molecule type: protein
A;Residues: 28-66 <BOH2>
R;Couch, J.L.; Betley, M.J.
J. Bacteriol. 171, 4507-4510, 1989
A;Title: Nucleotide sequence of the type C-3 staphylococcal enterotoxin gene suggests tha
A;Reference number: A33866; MUID:89327174; PMID:2473979
A;Accession: A33866
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-266 <COU>
A;Cross-references: GB:M28364; NID:g153003; PIDN:AAA26624.1; PID:g153004
C;Genetics:
A;Gene: entC2
C;Superfamily: enterotoxin B
F;1-27/Domain: signal sequence #status predicted <SIG>
F;28-266/Product: enterotoxin C-2 #status predicted <MAT>

Query Match 46.1%; Score 543; DB 2; Length 266;
Best Local Similarity 46.4%; Pred. No. 9.8e-35;

Matches 109; Conservative 45; Mismatches 67; Indels 14; Gaps 6;

```
Qy          1 QQDPDPSQLHRSS-LVKNLQNIYFLYEGDPVTHENVKSVDQLLSHHLIYNVSG---PNYD 56
              | || | :||:||     : |: :||:   |:   | |||: |:| ||||:|     |||
Db         30 QPDPTPDELHKSSEFTGTMGNMKYLYDDHYVSATKVMSVDKFLAHDLIYNISDKKLKNYD 89

Qy         57 KLKTELKNQEMATLFKDKNVDIYGVEYYHLCYLC--ENAER----SACIYGGVTNHEGNH 110
              |:|||| |:::|  :||: ||:||  ||  ||     :|  :       |:|||:| |||||
Db         90 KVKTELLNEDLAKKYKDEVVDVYGSNYYVNCYFSSKDNVGKVTGGKTCMYGGITKHEGNH 149

Qy        111 LEIP--KKIVVKVSIDGIQSLSFDIETNKKMVTAQELDYKVRKYLTDNKQLYTNGPSKYE 168
               :    : ::::|  :   ::||:::|:|| |||||| | | :| : | ||    | ||
Db        150 FDNGNLQNVLIRVYENKRNTISFEVQTDKKSVTAQELDIKARNFLINKKNLYEFNSSPYE 209

Qy        169 TGYIKFIPKNKESFWFDFFPEP--EFTQSKYLMIYKDNETLDSNTSQIEVYLTTK 221
              |||||||  |  :||:|  | |  :| ||||||:| ||:|:|| : :|||:||||
Db        210 TGYIKFIENNGNTFWYDMMPAPGDKFDQSKYLMMYNDNKTVDSKSVKIEVHLTTK 264
```

RESULT 9
ENSAC1
enterotoxin C-1 precursor - Staphylococcus aureus
C;Species: Staphylococcus aureus
C;Date: 15-Nov-1984 #sequence_revision 05-Jan-1996 #text_change 18-Jun-1999
C;Accession: S06356; A01816
R;Bohach, G.A.; Schlievert, P.M.
Mol. Gen. Genet. 209, 15-20, 1987
A;Title: Nucleotide sequence of the staphylococcal enterotoxin C1 gene and relatedness t
A;Reference number: S06356; MUID:88038352; PMID:2823067
A;Accession: S06356
A;Molecule type: DNA
A;Residues: 1-266 <BOH>
A;Cross-references: EMBL:X05815; NID:g46566; PIDN:CAA29260.1; PID:g46567
R;Schmidt, J.J.; Spero, L.
J. Biol. Chem. 258, 6300-6306, 1983
A;Title: The complete amino acid sequence of Staphylococcal enterotoxin C-1.
A;Reference number: A01816; MUID:83213327; PMID:6189824
A;Accession: A01816
A;Molecule type: protein
A;Residues: 28-75,'IL',78-176,'N',178-266 <SCH>
C;Genetics:
A;Gene: entC1
C;Superfamily: enterotoxin B
C;Keywords: enterotoxin
F;1-27/Domain: signal sequence #status predicted <SIG>
F;28-266/Product: enterotoxin C-1 #status experimental <MAT>
F;120-137/Disulfide bonds: #status experimental

Query Match 46.0%; Score 542; DB 1; Length 266;
Best Local Similarity 46.8%; Pred. No. 1.2e-34;
Matches 110; Conservative 44; Mismatches 67; Indels 14; Gaps 6;

```
Qy          1 QQDPDPSQLHRSSLVKNL-QNIYFLYEGDPVTHENVKSVDQLLSHHLIYNVSG---PNYD 56
              | || | :||::|    | :|:  ||:   |:   |||||: |:| ||||:|     |||
Db         30 QPDPTPDELHKASKFTGLMENMKVLYDDHYVSATKVKSVDKFLAHDLIYNISDKKLKNYD 89

Qy         57 KLKTELKNQEMATLFKDKNVDIYGVEYYHLCYLC--ENAER----SACIYGGVTNHEGNH 110
              |:|||| |: :|  :||: ||:||  ||  ||     :|  :       |:|||:| |||||
Db         90 KVKTELLNEGLAKKYKDEVVDVYGSNYYVNCYFSSKDNVGKVTGGKTCMYGGITKHEGNH 149

Qy        111 LEIP--KKIVVKVSIDGIQSLSFDIETNKKMVTAQELDYKVRKYLTDNKQLYTNGPSKYE 168
               :    : ::::|  :   ::||:::|:|| |||||| | | :| : | ||    | ||
Db        150 FDNGNLQNVLIRVYENKRNTISFEVQTDKKSVTAQELDIKARNFLINKKNLYEFNSSPYE 209

Qy        169 TGYIKFIPKNKESFWFDFFPEP--EFTQSKYLMIYKDNETLDSNTSQIEVYLTTK 221
              |||||||  |  :||:|  | |  :| ||||||:| ||:|:|| : :|||:||||
Db        210 TGYIKFIENNGNTFWYDMMPAPGDKFDQSKYLMMYNDNKTVDSKSVKIEVHLTTK 264
```

RESULT 10
G89968
extracellular enterotoxin type G precursor [imported] - Staphylococcus aureus (strain N3
C;Species: Staphylococcus aureus
C;Date: 10-May-2001 #sequence_revision 10-May-2001 #text_change 22-Oct-2001
C;Accession: G89968
R;Kuroda, M.; Ohta, T.; Uchiyama, I.; Baba, T.; Yuzawa, H.; Kobayashi, I.; Cui, L.; Oguc
ma, A.; Mizutani-Ui, Y.; Kobayashi, N.; Sawano, T.; Inoue, R.; Kaito, C.; Sekimizu, K.;
C.; Shiba, T.; Hattori, M.; Ogasawara, N.; Hayashi, H.; Hiramatsu, K.
Lancet 357, 1225-1240, 2001
A;Title: Whole genome sequencing of meticillin-resistant Stapylococcus aureus.
A;Reference number: A89758; MUID:21311952; PMID:11418146
A;Accession: G89968
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-258 <KUR>
A;Cross-references: GB:BA000018; PID:g13701617; PIDN:BAB42910.1; GSPDB:GN00149
A;Experimental source: strain N315
C;Genetics:
A;Gene: seg
C;Superfamily: enterotoxin B

Query Match 39.7%; Score 468; DB 2; Length 258;
Best Local Similarity 43.8%; Pred. No. 5.8e-29;
Matches 102; Conservative 39; Mismatches 78; Indels 14; Gaps 4;

```
Qy          1 QQDPDPSQLHRSSLVKN----LQNIYFLYEGDPVTHENVKSVDQLLSHHLIYNVSGPNYD 56
              | ||   :|:: |  ||    : |:  ||   ||    | :  | ||| ||: :   :|:
Db         26 QPDPKLDELNKVSDYKNNKGTMGNVMNLYTSPPVEGRGVINSRQFLSHDLIFPIEYKSYN 85

Qy         57 KLKTELKNQEMATLFKDKNVDIYGVEYYHLCYLCENAE------RSACIYGGVTNHEGNH 110
              ::|||:| |:|  :||| |||:|| |:: | : ::           |:|||:| :   :
Db         86 EVKTELENTELANNYKDKKVDIFGVPYFYTCIIPKSEPDINQNFGGCCMYGGLTFNSSEN 145

Qy        111 LEIPKKIVVKVSIDGIQSLSFDIETNKKMVTAQELDYKVRKYLTDNKQLYTNGPSKYETG 170
               | | | |:|:||  ||| | | ||| ||| |||||| | :||  |:||    | :|:|
Db        146 -ERDKLITVQVTIDNRQSLGFTITTNKNMVTIQELDYKARHWLTKEKKLYEFDGSAFESG 204

Qy        171 YIKFIPKNKESFWFDFFPEPE---FTQSKYLMIYKDNETLDSNTSQIEVYLTT 220
              ||||  ||  ||||| ||:  |   |  |:| || ||: :|| : ::||:| |
Db        205 YIKFTEKNNTSFWFDLFPKKELVPFVPYKFLNIYGDNKVVDSKSIKMEVFLNT 257
```

RESULT 11
A89969
enterotoxin YENT2 [imported] - Staphylococcus aureus (strain N315)
C;Species: Staphylococcus aureus
C;Date: 10-May-2001 #sequence_revision 10-May-2001 #text_change 22-Oct-2001
C;Accession: A89969
R;Kuroda, M.; Ohta, T.; Uchiyama, I.; Baba, T.; Yuzawa, H.; Kobayashi, I.; Cui, L.; Oguc
ma, A.; Mizutani-Ui, Y.; Kobayashi, N.; Sawano, T.; Inoue, R.; Kaito, C.; Sekimizu, K.;
C.; Shiba, T.; Hattori, M.; Ogasawara, N.; Hayashi, H.; Hiramatsu, K.
Lancet 357, 1225-1240, 2001
A;Title: Whole genome sequencing of meticillin-resistant Stapylococcus aureus.
A;Reference number: A89758; MUID:21311952; PMID:11418146
A;Accession: A89969
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-136 <KUR>
A;Cross-references: GB:BA000018; PID:g13701619; PIDN:BAB42912.1; GSPDB:GN00149
A;Experimental source: strain N315
C;Genetics:
A;Gene: yent2

Query Match 28.6%; Score 337; DB 2; Length 136;
Best Local Similarity 49.6%; Pred. No. 3.4e-19;
Matches 67; Conservative 22; Mismatches 38; Indels 8; Gaps 2;

```
Qy         95 RSACIYGGVTNHEGNHL------EIPKKIVVKVSIDGIQSLSFDIETNKKMVTAQELDYK 148
              :  |:||||| |:|| :       :    |::||  :   |||||| |||| :||||:|||
Db          2 KKTCMYGGVTEHDGNQIDKNNSTDNSHNILIKVYENERNSLSFDIPTNKKNITAQEIDYK 61

Qy        149 VRKYLTDNKQLYTNGPSKYETGYIKFIPKNKESFWFDFFPE--PEFTQSKYLMIYKDNET 206
              || ||  :| ||     | |||||||||  :  |||:|  ||   :|  :|||:|| ||:|
Db         62 VRNYLLKHKNLYEFNSSPYETGYIKFIEGSGHSFWYDLMPESGKKFYPTKYLLIYNDNKT 121

Qy        207 LDSNTSQIEVYLTTK 221
```

```
Db        122 VESKSINVEVHLTKK 136
```

RESULT 12
A33953
enterotoxin D precursor - Staphylococcus aureus
C;Species: Staphylococcus aureus
C;Date: 09-Mar-1990 #sequence_revision 09-Mar-1990 #text_change 15-Oct-1999
C;Accession: A33953
R;Bayles, K.W.; Iandolo, J.J.
J. Bacteriol. 171, 4799-4806, 1989
A;Title: Genetic and molecular analyses of the gene encoding staphylococcal enterotoxin
A;Reference number: A33953; MUID:89359112; PMID:2549000
A;Accession: A33953
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-258 <BAY>
A;Cross-references: GB:M28521; NID:g1492109; PIDN:AAB06195.1; PID:g758691
C;Superfamily: enterotoxin B

  Query Match   28.3%; Score 333.5; DB 2; Length 258;
  Best Local Similarity   36.7%; Pred. No. 1.4e-18;
  Matches   83; Conservative   42; Mismatches   86; Indels   15; Gaps   7;

```
Qy          8 QLHR-----SSLVKNLQNIYFLYEGDPVTHENVKSVDQLLSHHLIYN---VSGPNYDKLK 59
              :||:     |: : |::: |   : :|:  ||  : || | : |:|       |:: |
Db         36 ELHKKSELSSSTALNNMKHSY--ADKNPIIGENKSTGDQFLENTLLYKKFFTDLIPFEDLL 93

Qy         60 TELKNQEMATLFKDKNVDIYGVEYYHLCYLCENAERSACIYGGVTNHEGNHLEIPKKIVV 119
                ::|||  || ||||:| : |    ||  |  :|:|| ||||| |||| |:  ||| :
Db         94 INFNSKEMAQHFKSKNVDVYPIRYSINCYGGE-IDRTACTYGGVTPHEGNKLKERKKIPI 152

Qy        120 KVSIDGIQ-SLSFD-IETNKKMVTAQELDYKVRKYLTDNKQLYTNGP--SKYETGYIKFI 175
               : |:|:|  :| | ::|:|| || |||| : |:||  : :|| |      | : | |:|
Db        153 NLWINGVQKEVSLDKVQTDKKNVTVQELDAQARRYLQKDLKLYNNDTLGGKIQRGKIEFD 212

Qy        176 PKNKESFWFDFFPEPEFTQSKYLMIYKDNETLDSNTSQIEVYLTTK 221
                :     :| |          | | || ||:|| :    |::||  |
Db        213 SSDGSKVSYDLFDVKGDFPEKQLRIYSDNKTLSTEHLHIDIYLYEK 258
```

RESULT 13
E89969
enterotoxin SeO [imported] - Staphylococcus aureus (strain N315)
C;Species: Staphylococcus aureus
C;Date: 10-May-2001 #sequence_revision 10-May-2001 #text_change 22-Oct-2001
C;Accession: E89969
R;Kuroda, M.; Ohta, T.; Uchiyama, I.; Baba, T.; Yuzawa, H.; Kobayashi, I.; Cui, L.; Oguc
ma, A.; Mizutani-Ui, Y.; Kobayashi, N.; Sawano, T.; Inoue, R.; Kaito, C.; Sekimizu, K.;
 C.; Shiba, T.; Hattori, M.; Ogasawara, N.; Hayashi, H.; Hiramatsu, K.
Lancet 357, 1225-1240, 2001
A;Title: Whole genome sequencing of meticillin-resistant Stapylococcus aureus.
A;Reference number: A89758; MUID:21311952; PMID:11418146
A;Accession: E89969
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-260 <KUR>
A;Cross-references: GB:BA000018; PID:g13701623; PIDN:BAB42916.1; GSPDB:GN00149
A;Experimental source: strain N315
C;Genetics:
A;Gene: seo
C;Superfamily: enterotoxin B

  Query Match   27.5%; Score 323.5; DB 2; Length 260;
  Best Local Similarity   38.5%; Pred. No. 8.6e-18;
  Matches   90; Conservative   33; Mismatches   94; Indels   17; Gaps   7;

```
Qy          1 QQDPDPSQLHRSSLVK--NLQNIYFLYEGDPVTHENVKSV----DQLLSHHLIYN----V 50
              ::||     | : | |     | ||   |  :  |   |||:    :: |   |::     :
Db         30 EEDPKIESLCKKSSVDPIALHNINDDYINNRFT--TVKSIVSTTEKFLDFDLLFKSINWL 87

Qy         51 SG--PNYDKLKTELKNQEMATLFKDKNVDIYGVEYYHLCYLCENAERSACIYGGVTNHEG 108
```

```
Db         88 DGISAEFKDLKVEFSSSAISKEFLGKTVDIYGVYYKAHCH-GEHQVDTACTYGGVTPHEN 146

Qy        109 NHLEIPKKIVVKVSIDGIQSLSFDIETNKKMVTAQELDYKVRKYLTDNKQLYTNGPSKYE 168
              | |  || | | |  | :   :| : |:|| |||||||| |||  | :  :||     | :
Db        147 NKLSEPKNIGVAVYKDNVNVNTFIVTTDKKKVTAQELDIKVRTKLNNAYKLYDRMTSDVQ 206

Qy        169 TGYIKF--IPKNKESFWFDFFPEPEFTQSKYLMIYKDNETLDSNTSQIEVYLTT 220
              |||||    ::||||::| |        :|| || ||:|:||:    |:||| |
Db        207 KGYIKFHSHSEHKESFYYDLFYIKGNLPDQYLQIYNDNKTIDSSDYHIDVYLFT 260
```

RESULT 14
C89984
enterotoxin P [imported] - Staphylococcus aureus (strain N315)
C;Species: Staphylococcus aureus
C;Date: 10-May-2001 #sequence_revision 10-May-2001 #text_change 22-Oct-2001
C;Accession: C89984
R;Kuroda, M.; Ohta, T.; Uchiyama, I.; Baba, T.; Yuzawa, H.; Kobayashi, I.; Cui, L.; Oguch
ma, A.; Mizutani-Ui, Y.; Kobayashi, N.; Sawano, T.; Inoue, R.; Kaito, C.; Sekimizu, K.; H
 C.; Shiba, T.; Hattori, M.; Ogasawara, N.; Hayashi, H.; Hiramatsu, K.
Lancet 357, 1225-1240, 2001
A;Title: Whole genome sequencing of meticillin-resistant Stapylococcus aureus.
A;Reference number: A89758; MUID:21311952; PMID:11418146
A;Accession: C89984
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-260 <KUR>
A;Cross-references: GB:BA000018; PID:g13701743; PIDN:BAB43036.1; GSPDB:GN00149
A;Experimental source: strain N315
C;Genetics:
A;Gene: sep
C;Superfamily: enterotoxin B

  Query Match   26.8%; Score 315.5; DB 2; Length 260;
  Best Local Similarity   36.7%; Pred. No. 3.6e-17;
  Matches   84; Conservative   44; Mismatches   74; Indels   27; Gaps   12;

```
Qy          7 SQLHRSSLVKNLQNIYFLYEGDPVTHENVKSVDQLLSHHLIYN--VSGPN-YDKLKTELK 63
              |:|  ::|  ||:  |: : |  :  || :| || | : :::|   :|   |: |  :|
Db         43 SELQGTAL-SNLRQTYY-HNGSAII-ENKESNDQFLKNTILFNDFFTGHQWYNDLLVDLG 99

Qy         64 NQEMATLFKDKNVDIYGVEYYHLCYLCENAE--RSACIYGGVTNHEGNHLEIPKKIVVKV 121
              ::: | ::| | ||:||| |     | |       ::||:||||| |: | ||  ||: : :
Db        100 SKDTANIYKGKKVDLYGVYY---GYQCTGGTPFKTACMYGGVTLHDNNQLEEEKKVPINL 156

Qy        122 SIDGIQSL--SFDIETNKKMVTAQELDYKVRKYLTDNKQLYT----NGPSKYETGYIKFI 175
               ||| |:      ::|||| || |||| : | || :   ||      ||  | : | |:|
Db        157 WIDGKQNTVPLGTVKTNKKEVTVQELDLQSRHYLHETYNLYNTDAFNG--KIQRGLIEFH 214

Qy        176 PKNKESFWFDFFPEPEFTQSKY----LMIYKDNETLDSNTSQIEVYLTT 220
              | : :|  :| |     | :|     | ||:||:|: |     |::|| |
Db        215 PSSGDSVGYDLFG----AQGQYPDTQLRIYRDNKTIKSKNMHIDIYLYT 259
```

RESULT 15
H89968
enterotoxin SeN [imported] - Staphylococcus aureus (strain N315)
C;Species: Staphylococcus aureus
C;Date: 10-May-2001 #sequence_revision 10-May-2001 #text_change 22-Oct-2001
C;Accession: H89968
R;Kuroda, M.; Ohta, T.; Uchiyama, I.; Baba, T.; Yuzawa, H.; Kobayashi, I.; Cui, L.; Oguch
ma, A.; Mizutani-Ui, Y.; Kobayashi, N.; Sawano, T.; Inoue, R.; Kaito, C.; Sekimizu, K.; H
 C.; Shiba, T.; Hattori, M.; Ogasawara, N.; Hayashi, H.; Hiramatsu, K.
Lancet 357, 1225-1240, 2001
A;Title: Whole genome sequencing of meticillin-resistant Stapylococcus aureus.
A;Reference number: A89758; MUID:21311952; PMID:11418146
A;Accession: H89968
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-258 <KUR>
A;Cross-references: GB:BA000018; PID:g13701618; PIDN:BAB42911.1; GSPDB:GN00149
A;Experimental source: strain N315

C;Genetics:
A;Gene: sen
C;Superfamily: enterotoxin B

Query Match 25.3%; Score 297.5; DB 2; Length 258;
Best Local Similarity 35.8%; Pred. No. 8.6e-16;
Matches 77; Conservative 39; Mismatches 86; Indels 13; Gaps 6;

```
Qy          12 SSLVKNLQNIY--FLYEGDPVTHENVKSVDQLLSHHLI---YNVSGPNYDKLKTELKNQE 66
               || : || : |    :: |     |  | ||||:: :|    ::|       || |  : :
Db          46 SSKLFNLTSYYTDITWQLD---ESNKISTDQLLNNTIILKNIDISVLKTSSLKVEFNSSD 102

Qy          67 MATLFKDKNVDIYGVEYYHLCY-LCENAERSACIYGGVTNHEGNHLEIPKKIVVKVSIDG 125
               :|  || ||:||||: : : |  | |  |:::|:|||| |:|| |:  | | | |  ||
Db         103 LANQFKGKNIDIYGLYFGNKCVGLTE--EKTSCLYGGVTIHDGNQLDEEKVIGVNVFKDG 160

Qy         126 IQSLSFDIETNKKMVTAQELDYKVRKYLTDNKQLYTNGPSKYETGYIKFIPKN--KESFW 183
               :|   | |:| |  || |||| |||  | :  ::|       : | | |   |   :||:
Db         161 VQQEGFVIKTKKAKVTVQELDTKVRFKLENLYKIYNKDTGNIQKGCIFFHSHNHQDQSFY 220

Qy         184 FDFFPEPEFTQSKYLMIYKDNETLDSNTSQIEVYL 218
               :| :       :::   | || |: |:   |:|:|
Db         221 YDLYNVKGSVGAEFFQFYSDNRTVSSSNYHIDVFL 255
```

Search completed: May 7, 2004, 12:09:40
Job time : 22 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: May 7, 2004, 12:01:27 ; Search time 17 Seconds
(without alignments)
676.912 Million cell updates/sec

Title: US-10-625-221-14
Perfect score: 1178
Sequence: 1 QQDPDPSQLHRSSLVKNLQN.....KDNETLDSNTSQIEVYLTTK 221

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 141681 seqs, 52070155 residues

Total number of hits satisfying chosen parameters: 141681

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : SwissProt_42:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	1169	99.2	251	1	SPEA_STRPY	P08095 streptococc
2	572	48.6	266	1	ETXB_STAAU	P01552 staphylococ
3	546	46.3	266	1	ETC3_STAAM	P23313 staphylococ
4	543	46.1	266	1	ETC2_STAAU	P34071 staphylococ
5	542	46.0	266	1	ETC1_STAAU	P01553 staphylococ
6	468	39.7	258	1	ETXG_STAAM	O85382 staphylococ
7	333.5	28.3	258	1	ETXD_STAAU	P20723 staphylococ
8	293.5	24.9	257	1	ETXA_STAAW	P13163 staphylococ
9	291.5	24.7	257	1	ETXE_STAAU	P12993 staphylococ
10	281.5	23.9	241	1	ETXH_STAAW	Q53585 staphylococ
11	209.5	17.8	236	1	SPEH_STRPY	Q9x5c8 streptococc
12	205	17.4	235	1	SPEC_STRPY	P13380 streptococc
13	162	13.8	234	1	SPEG_STRPY	Q9x5c7 streptococc
14	101	8.6	234	1	TSST_STAAU	P06886 staphylococ
15	101	8.6	894	1	DSC1_HUMAN	Q08554 homo sapien
16	96	8.1	396	1	PRRC_ECOLI	P17223 escherichia
17	96	8.1	908	1	DPO1_BORBU	O51498 borrelia bu
18	95.5	8.1	999	1	HGP2_HAEIN	P44809 haemophilus
19	93	7.9	434	1	PURA_WIGBR	Q8d322 wiggleswort
20	92	7.8	660	1	PRIA_BORBU	Q45032 borrelia bu
21	91	7.7	384	1	SCS7_YEAST	Q03529 saccharomyc
22	91	7.7	1956	1	ATX1_PLAFA	Q04956 plasmodium
23	90.5	7.7	3255	1	POLG_LMVE	P89876 l genome po
24	90	7.6	554	1	Y478_RICPR	Q9zd66 rickettsia
25	90	7.6	598	1	PRIM_STAEP	Q8cp23 staphylococ
26	90	7.6	3255	1	POLG_LMV0	P31999 l genome po
27	89.5	7.6	439	1	NADB_CLOPE	Q8xne2 clostridium
28	89	7.6	968	1	RNE_BUCAP	Q8k9j9 buchnera ap
29	88.5	7.5	495	1	Y243_AQUAE	O66606 aquifex aeo
30	88.5	7.5	540	1	MTA1_ACICA	P25201 acinetobact
31	88	7.5	181	1	RM06_ACACA	P46765 acanthamoeb
32	87.5	7.4	500	1	TACY_CLOPE	P19995 clostridium
33	87.5	7.4	2410	1	POL1_BAYMJ	Q01206 barley yell
34	87	7.4	322	1	GPI8_CAEEL	P49048 caenorhabdi
35	87	7.4	369	1	FIE_ARATH	Q9lt47 arabidopsis
36	87	7.4	492	1	MUE2_OCEIH	Q8cze4 oceanobacil
37	87	7.4	724	1	P85A_MOUSE	P26450 mus musculu
38	87	7.4	970	1	Y277_MYCGE	Q49409 mycoplasma
39	86.5	7.3	539	1	YJX3_YEAST	P47084 saccharomyc
40	86.5	7.3	625	1	RSC4_YEAST	Q02206 saccharomyc
41	86.5	7.3	758	1	PEPX_STRMU	Q8dvs2 streptococc
42	86.5	7.3	893	1	DSC1_BOVIN	Q01107 bos taurus
43	86.5	7.3	993	1	HGBC_HAEIN	Q9kiv0 haemophilus
44	86	7.3	1109	1	POL_CAEVC	P33459 caprine art
45	85.5	7.3	2412	1	POL1_BAYMG	Q04574 barley yell

ALIGNMENTS

```
RESULT 1
SPEA_STRPY
ID   SPEA_STRPY     STANDARD;      PRT;   251 AA.
AC   P08095;
DT   01-AUG-1988 (Rel. 08, Created)
DT   01-JAN-1990 (Rel. 13, Last sequence update)
DT   15-MAR-2004 (Rel. 43, Last annotation update)
DE   Exotoxin type A precursor (Scarlet fever toxin) (Erythrogenic toxin)
DE   (SPE A).
GN   SPEA OR SPYM18_0393.
OS   Streptococcus pyogenes, and
OS   Streptococcus pyogenes (serotype M18).
OC   Bacteria; Firmicutes; Lactobacillales; Streptococcaceae;
OC   Streptococcus.
OX   NCBI_TaxID=1314, 186103;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=86166804; PubMed=3514452;
RA   Weeks C.R., Ferretti J.J.;
RT   "Nucleotide sequence of the type A streptococcal exotoxin
RT   (erythrogenic toxin) gene from Streptococcus pyogenes bacteriophage
RT   T12.";
RL   Infect. Immun. 52:144-150(1986).
RN   [2]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=86284313; PubMed=3526093;
RA   Johnson L.P., L'Italien J.J., Schlievert P.M.;
RT   "Streptococcal pyrogenic exotoxin type A (scarlet fever toxin) is
RT   related to Staphylococcus aureus enterotoxin B.";
RL   Mol. Gen. Genet. 203:354-356(1986).
RN   [3]
RP   SEQUENCE FROM N.A.
RC   STRAIN=MGAS8232 / Serotype M18;
RX   MEDLINE=21927593; PubMed=11917108;
RA   Smoot J.C., Barbian K.D., Van Gompel J.J., Smoot L.M., Chaussee M.S.,
RA   Sylva G.L., Sturdevant D.E., Ricklefs S.M., Porcella S.F.,
RA   Parkins L.D., Beres S.B., Campbell D.S., Smith T.M., Zhang Q.,
RA   Kapur V., Daly J.A., Veasy L.G., Musser J.M.;
RT   "Genome sequence and comparative microarray analysis of serotype M18
RT   group A Streptococcus strains associated with acute rheumatic fever
RT   outbreaks.";
RL   Proc. Natl. Acad. Sci. U.S.A. 99:4668-4673(2002).
RN   [4]
RP   X-RAY CRYSTALLOGRAPHY (2.57 ANGSTROMS).
RX   MEDLINE=99094887; PubMed=9878045;
RA   Papageorgiou A.C., Collins C.M., Gutman D.M., Kline J.B.,
RA   O'Brien S.M., Tranter H.S., Acharya K.R.;
RT   "Structural basis for the recognition of superantigen streptococcal
RT   pyrogenic exotoxin A (SpeA1) by MHC class II molecules and T-cell
RT   receptors.";
RL   EMBO J. 18:9-21(1999).
CC   -!- FUNCTION: Causative agent of the symptoms associated with scarlet
CC       fever, have been associated with streptococcal toxic shock-like
CC       disease and may play a role in the early events of rheumatic
CC       fever.
CC   -!- SUBUNIT: Binds to major histocompatibility complex class II beta
```

```
CC       chain.
CC   -!- MISCELLANEOUS: THIS TOXIN SEEMS TO BE CODED BY BACTERIOPHAGE T12.
CC   -!- SIMILARITY: Belongs to the staphylococcal/streptococcal toxin
CC       family.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; U40453; AAC48868.1; -.
DR   EMBL; X03929; CAA27568.1; -.
DR   EMBL; AE009982; AAL97141.1; -.
DR   PIR; A26152; A26152.
DR   PDB; 1B1Z; 24-NOV-99.
DR   PDB; 1FNU; 17-NOV-00.
DR   PDB; 1FNV; 17-NOV-00.
DR   PDB; 1FNW; 17-NOV-00.
DR   PDB; 1HA5; 03-APR-02.
DR   PDB; 1L0X; 03-APR-02.
DR   InterPro; IPR008992; Bact_endotox.
DR   InterPro; IPR006177; Bctrl_tox.
DR   InterPro; IPR006123; Stap/Strep_toxin.
DR   InterPro; IPR006126; Staph/Strept_tox.
DR   InterPro; IPR006173; Staph_tox_OB.
DR   Pfam; PF02876; Stap_Strp_tox_C; 1.
DR   Pfam; PF01123; Stap_Strp_toxin; 1.
DR   PRINTS; PR00279; BACTRLTOXIN.
DR   PROSITE; PS00277; STAPH_STREP_TOXIN_1; 1.
DR   PROSITE; PS00278; STAPH_STREP_TOXIN_2; 1.
KW   Toxin; Signal; 3D-structure; Complete proteome.
FT   SIGNAL        1     30
FT   CHAIN        31    251       EXOTOXIN TYPE A.
FT   DISULFID    117    128
FT   CONFLICT      6      6       K -> E (IN REF. 2).
FT   CONFLICT     17     18       VT -> MK (IN REF. 2).
FT   CONFLICT     25     35       SQEVFAQQDPD -> LPKGICSTRPK (IN REF. 2).
FT   CONFLICT     40     40       H -> Q (IN REF. 2).
FT   CONFLICT     43     43       S -> N (IN REF. 2).
FT   CONFLICT     47     59       NLQNIYFLYEGDP -> TFKIYIFFMRVTL (IN
FT                                REF. 2).
FT   CONFLICT    129    129       I -> L (IN REF. 2).
FT   CONFLICT    165    178       TNKKMVTAQELDYK -> QIKNGNCSRISYT (IN
FT                                REF. 2).
FT   HELIX        36     38
FT   HELIX        42     44
FT   TURN         48     48
FT   HELIX        49     56
FT   STRAND       60     66
FT   STRAND       69     69
FT   TURN         73     74
FT   STRAND       75     78
FT   STRAND       82     82
FT   TURN         83     84
FT   STRAND       85     85
FT   STRAND       87     91
FT   HELIX        95    101
FT   TURN        102    103
FT   STRAND      105    110
FT   STRAND      113    113
FT   TURN        115    116
FT   STRAND      126    130
FT   STRAND      133    135
FT   TURN        137    138
FT   STRAND      140    153
FT   TURN        154    155
FT   STRAND      156    167
FT   STRAND      169    171
FT   HELIX       172    187
FT   STRAND      199    205
FT   STRAND      212    215
FT   HELIX       224    227
FT   HELIX       228    231
FT   TURN        232    233
FT   STRAND      236    238
FT   TURN        239    241
FT   STRAND      243    249
SQ   SEQUENCE   251 AA;  29246 MW;  54001FE4CCCBFCC3 CRC64;

  Query Match             99.2%;  Score 1169;  DB 1;  Length 251;
  Best Local Similarity   99.5%;  Pred. No. 2.4e-82;
  Matches  220;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy          1 QQDPDPSQLHRSSLVKNLQNIYFLYEGDPVTHENVKSVDQLLSHHLIYNVSGPNYDKLKT 60
              ||||||||||||||||||||||||||||||||||||||||||| ||||||||||||||||
Db         31 QQDPDPSQLHRSSLVKNLQNIYFLYEGDPVTHENVKSVDQLLSHDLIYNVSGPNYDKLKT 90

Qy         61 ELKNQEMATLFKDKNVDIYGVEYYHLCYLCENAERSACIYGGVTNHEGNHLEIPKKIVVK 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         91 ELKNQEMATLFKDKNVDIYGVEYYHLCYLCENAERSACIYGGVTNHEGNHLEIPKKIVVK 150

Qy        121 VSIDGIQSLSFDIETNKKMVTAQELDYKVRKYLTDNKQLYTNGPSKYETGYIKFIPKNKE 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        151 VSIDGIQSLSFDIETNKKMVTAQELDYKVRKYLTDNKQLYTNGPSKYETGYIKFIPKNKE 210

Qy        181 SFWFDFFPEPEFTQSKYLMIYKDNETLDSNTSQIEVYLTTK 221
              |||||||||||||||||||||||||||||||||||||||||
Db        211 SFWFDFFPEPEFTQSKYLMIYKDNETLDSNTSQIEVYLTTK 251


RESULT 2
ETXB_STAAU
ID   ETXB_STAAU     STANDARD;      PRT;   266 AA.
AC   P01552;
DT   21-JUL-1986 (Rel. 01, Created)
DT   13-AUG-1987 (Rel. 05, Last sequence update)
DT   15-MAR-2004 (Rel. 43, Last annotation update)
DE   Enterotoxin type B precursor (SEB).
GN   ENTB.
OS   Staphylococcus aureus.
OC   Bacteria; Firmicutes; Bacillales; Staphylococcus.
OX   NCBI_TaxID=1280;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=86168029; PubMed=3957869;
RA   Jones C.L., Khan S.A.;
RT   "Nucleotide sequence of the enterotoxin B gene from Staphylococcus
RT   aureus.";
RL   J. Bacteriol. 166:29-33(1986).
RN   [2]
RP   SEQUENCE OF 40-91 FROM N.A.
RX   MEDLINE=85298255; PubMed=3898073;
RA   Ranelli D.M., Jones C.L., Johns M.B., Mussey G.J., Khan S.A.;
RT   "Molecular cloning of staphylococcal enterotoxin B gene in
RT   Escherichia coli and Staphylococcus aureus.";
RL   Proc. Natl. Acad. Sci. U.S.A. 82:5850-5854(1985).
RN   [3]
RP   SEQUENCE OF 28-266 (S-6).
RX   MEDLINE=71007902; PubMed=5470821;
RA   Huang I.-Y., Bergdoll M.S.;
RT   "The primary structure of staphylococcal enterotoxin B. 3. The
RT   cyanogen bromide peptides of reduced and aminoethylated enterotoxin
RT   B, and the complete amino acid sequence.";
RL   J. Biol. Chem. 245:3518-3525(1970).
RN   [4]
RP   X-RAY CRYSTALLOGRAPHY (2.5 ANGSTROMS).
RX   MEDLINE=93063291; PubMed=1436058;
RA   Swaminathan S., Furey W.F. Jr., Pletcher J., Sax M.;
RT   "Crystal structure of staphylococcal enterotoxin B, a superantigen.";
RL   Nature 359:801-806(1992).
RN   [5]
RP   X-RAY CRYSTALLOGRAPHY (2.7 ANGSTROMS) OF COMPLEX WITH MHC II.
RX   MEDLINE=94203282; PubMed=8152483;
```

```
RA   Jardetzky T.S., Brown J.H., Gorga J.C., Stern L.J., Urban R.G.,
RA   Chi Y.I., Stauffacher C., Strominger J.L., Wiley D.C.;
RT   "Three-dimensional structure of a human class II histocompatibility
RT   molecule complexed with superantigen.";
RL   Nature 368:711-718(1994).
RN   [6]
RP   X-RAY CRYSTALLOGRAPHY (2.4 ANGSTROMS) OF COMPLEX WITH TCR.
RX   MEDLINE=99096298; PubMed=9881971;
RA   Li H., Llera A., Tsuchiya D., Leder L., Ysern X., Schlievert P.M.,
RA   Karjalainen K., Mariuzza R.A.;
RT   "Three-dimensional structure of the complex between a T cell receptor
RT   beta chain and the superantigen staphylococcal enterotoxin B.";
RL   Immunity 9:807-816(1998).
RN   [7]
RP   X-RAY CRYSTALLOGRAPHY (1.5 ANGSTROMS).
RX   MEDLINE=98181012; PubMed=9514739;
RA   Papageorgiou A.C., Tranter H.S., Acharya K.R.;
RT   "Crystal structure of microbial superantigen staphylococcal
RT   enterotoxin B at 1.5-A resolution: implications for superantigen
RT   recognition by MHC class II molecules and T-cell receptors.";
RL   J. Mol. Biol. 277:61-79(1998).
CC   -!- FUNCTION: Staphylococcal enterotoxins cause the intoxication
CC       staphylococcal food poisoning syndrome. The illness characterized
CC       by high fever, hypotension, diarrhea, shock, and in some cases
CC       death.
CC   -!- SUBCELLULAR LOCATION: Secreted.
CC   -!- SIMILARITY: Belongs to the staphylococcal/streptococcal toxin
CC       family.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; M11118; AAA88550.1; -.
DR   PIR; S27360; ENSAB6.
DR   PDB; 1SEB; 20-JUN-96.
DR   PDB; 2SEB; 28-JAN-98.
DR   PDB; 3SEB; 27-MAY-98.
DR   PDB; 1SE3; 16-JUN-97.
DR   PDB; 1SE4; 15-OCT-97.
DR   PDB; 1SBB; 04-MAR-99.
DR   PDB; 1D5M; 14-MAR-01.
DR   PDB; 1D5X; 28-JUN-00.
DR   PDB; 1D5Z; 28-JUN-00.
DR   PDB; 1D6E; 28-JUN-00.
DR   InterPro; IPR008992; Bact_endotox.
DR   InterPro; IPR006177; Bctrl_tox.
DR   InterPro; IPR006123; Stap/Strep_toxin.
DR   InterPro; IPR006126; Staph/Strept_tox.
DR   InterPro; IPR006173; Staph_tox_OB.
DR   Pfam; PF02876; Stap_Strp_tox_C; 1.
DR   Pfam; PF01123; Stap_Strp_toxin; 1.
DR   PRINTS; PR00279; BACTRLTOXIN.
DR   PROSITE; PS00277; STAPH_STREP_TOXIN_1; 1.
DR   PROSITE; PS00278; STAPH_STREP_TOXIN_2; 1.
KW   Enterotoxin; Toxin; Signal; Superantigen; 3D-structure.
FT   SIGNAL        1     27
FT   CHAIN        28    266       ENTEROTOXIN TYPE B.
FT   DISULFID    120    140
FT   CONFLICT     56     58       DDN -> NND (IN REF. 3).
FT   CONFLICT     69     77       DQFLYFDLI -> NEFFDLIYL (IN REF. 3).
FT   CONFLICT    118    118       MISSING (IN REF. 3).
FT   CONFLICT    128    130       DIN -> NID (IN REF. 3).
FT   CONFLICT    133    135       QTD -> ENT (IN REF. 3).
FT   CONFLICT    149    150       NG -> GN (IN REF. 3).
FT   CONFLICT    156    156       Y -> YY (IN REF. 3).
FT   CONFLICT    185    186       QE -> EQ (IN REF. 3).
FT   CONFLICT    233    233       D -> N (IN REF. 3).
FT   CONFLICT    246    247       DN -> ND (IN REF. 3).
FT   STRAND       29     29
FT   HELIX        41     43
FT   STRAND       44     44
FT   TURN         48     48
FT   HELIX        49     52
FT   HELIX        53     55
FT   STRAND       60     66
FT   STRAND       69     69
FT   TURN         73     74
FT   STRAND       75     78
FT   TURN         83     85
FT   STRAND       90     94
FT   HELIX        98    104
FT   TURN        105    106
FT   STRAND      108    113
FT   STRAND      116    116
FT   TURN        118    119
FT   TURN        127    128
FT   STRAND      138    142
FT   STRAND      145    147
FT   TURN        149    150
FT   STRAND      152    165
FT   TURN        166    167
FT   STRAND      168    179
FT   STRAND      181    183
FT   HELIX       184    199
FT   STRAND      209    218
FT   TURN        219    220
FT   STRAND      221    226
FT   STRAND      232    232
FT   HELIX       237    241
FT   HELIX       242    246
FT   STRAND      249    251
FT   TURN        252    254
FT   STRAND      255    263
SQ   SEQUENCE   266 AA;  31436 MW;  B6D417F61CF018B0 CRC64;

  Query Match             48.6%;  Score 572;  DB 1;  Length 266;
  Best Local Similarity   49.2%;  Pred. No. 9e-37;
  Matches  116;  Conservative   36;  Mismatches   68;  Indels   16;  Gaps    5;

Qy          1 QQDPDPSQLHRSSLVKNL-QNIYFLYEGDPVTHENVKSVDQLLSHHLIYNVSGP---NYD 56
              | || | :||:||     | :|:  ||: : |:  ||||:|| |   |||::       |||
Db         30 QPDPKPDELHKSSKFTGLMENMKVLYDDNHVSAINVKSIDQFLYFDLIYSIKDTKLGNYD 89

Qy         57 KLKTELKNQEMATLFKDKNVDIYGVEYYHLCYLCENA---------ERSACIYGGVTNHE 107
               :: | ||:::|  :||| ||::|  ||: ||  :            :|  |:|||| |
Db         90 NVRVEFKNKDLADKYKDKYVDVFGANYYYQCYFSKKTNDINSHQTDKRKTCMYGGVTEHN 149

Qy        108 GNHLEIPKKIVVKVSIDGIQSLSFDIETNKKMVTAQELDYKVRKYLTDNKQLYTNGPSKY 167
              || |:  : | |:|  ||   ||||::|||| |||||||  | ||  ||:||    | |
Db        150 GNQLDKYRSITVRVFEDGKNLLSFDVQTNKKKVTAQELDYLTRHYLVKNKKLYEFNNSPY 209

Qy        168 ETGYIKFIPKNKESFWFDFFPEP--EFTQSKYLMIYKDNETLDSNTSQIEVYLTTK 221
              |||||||| :|: |||:|  | |  :| |||||:| ||: :||   :||||||||
Db        210 ETGYIKFI-ENENSFWYDMMPAPGDKFDQSKYLMMYNDNKMVDSKDVKIEVYLTTK 264


RESULT 3
ETC3_STAAM
ID   ETC3_STAAM     STANDARD;      PRT;   266 AA.
AC   P23313;
DT   01-NOV-1991 (Rel. 20, Created)
DT   01-NOV-1991 (Rel. 20, Last sequence update)
DT   15-MAR-2004 (Rel. 43, Last annotation update)
DE   Enterotoxin type C-3 precursor (SEC3).
GN   ENTC3 OR SAV2009 OR SA1817.
OS   Staphylococcus aureus (strain Mu50 / ATCC 700699),
OS   Staphylococcus aureus (strain N315), and
OS   Staphylococcus aureus.
OC   Bacteria; Firmicutes; Bacillales; Staphylococcus.
OX   NCBI_TaxID=158878, 158879, 1280;
RN   [1]
```

```
RP   SEQUENCE FROM N.A.
RC   STRAIN=Mu50 / ATCC 700699, and N315;
RX   MEDLINE=21311952; PubMed=11418146;
RA   Kuroda M., Ohta T., Uchiyama I., Baba T., Yuzawa H., Kobayashi I.,
RA   Cui L., Oguchi A., Aoki K.-I., Nagai Y., Lian J.-Q., Ito T.,
RA   Kanamori M., Matsumaru H., Maruyama A., Murakami H., Hosoyama A.,
RA   Mizutani-Ui Y., Takahashi N.K., Sawano T., Inoue R.-I., Kaito C.,
RA   Sekimizu K., Hirakawa H., Kuhara S., Goto S., Yabuzaki J.,
RA   Kanehisa M., Yamashita A., Oshima K., Furuya K., Yoshino C., Shiba T.,
RA   Hattori M., Ogasawara N., Hayashi H., Hiramatsu K.;
RT   "Whole genome sequencing of meticillin-resistant Staphylococcus
RT   aureus.";
RL   Lancet 357:1225-1240(2001).
RN   [2]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=90220508; PubMed=2325627;
RA   Hovde C.J., Hackett S.P., Bohach G.A.;
RT   "Nucleotide sequence of the staphylococcal enterotoxin C3 gene:
RT   sequence comparison of all three type C staphylococcal
RT   enterotoxins.";
RL   Mol. Gen. Genet. 220:329-333(1990).
RN   [3]
RP   X-RAY CRYSTALLOGRAPHY (3.5 ANGSTROMS) OF COMPLEX WITH TCR.
RX   MEDLINE=97064178; PubMed=8906797;
RA   Fields B.A., Malchiodi E.L., Li H., Ysern X., Stauffacher C.V.,
RA   Schlievert P.M., Karjalainen K., Mariuzza R.A.;
RT   "Crystal structure of a T-cell receptor beta-chain complexed with a
RT   superantigen.";
RL   Nature 384:188-192(1996).
CC   -!- FUNCTION: Staphylococcal enterotoxins cause the intoxication
CC       staphylococcal food poisoning syndrome. The illness characterized
CC       by high fever, hypotension, diarrhea, shock, and in some cases
CC       death.
CC   -!- SUBCELLULAR LOCATION: Secreted.
CC   -!- SIMILARITY: Belongs to the staphylococcal/streptococcal toxin
CC       family.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; AP003364; BAB58171.1; -.
DR   EMBL; AP003135; BAB43097.1; -.
DR   EMBL; X51661; CAA35972.1; -.
DR   PIR; S11885; S11885.
DR   PDB; 1JCK; 12-NOV-97.
DR   PDB; 1KLG; 02-AUG-02.
DR   PDB; 1KLU; 14-AUG-02.
DR   InterPro; IPR008992; Bact_endotox.
DR   InterPro; IPR006177; Bctrl_tox.
DR   InterPro; IPR006123; Stap/Strep_toxin.
DR   InterPro; IPR006126; Staph/Strept_tox.
DR   InterPro; IPR006173; Staph_tox_OB.
DR   Pfam; PF02876; Stap_Strp_tox_C; 1.
DR   Pfam; PF01123; Stap_Strp_toxin; 1.
DR   PRINTS; PR00279; BACTRLTOXIN.
DR   PROSITE; PS00277; STAPH_STREP_TOXIN_1; 1.
DR   PROSITE; PS00278; STAPH_STREP_TOXIN_2; 1.
KW   Enterotoxin; Toxin; Signal; Superantigen; 3D-structure;
KW   Complete proteome.
FT   SIGNAL        1     27
FT   CHAIN        28    266       ENTEROTOXIN TYPE C-3.
FT   DISULFID    120    137
SQ   SEQUENCE   266 AA;  30671 MW;  5ED8A32D11FFCA59 CRC64;

  Query Match             46.3%;  Score 546;  DB 1;  Length 266;
  Best Local Similarity   47.2%;  Pred. No. 8.7e-35;
  Matches  111;  Conservative   43;  Mismatches   67;  Indels   14;  Gaps    6;

Qy          1 QQDPDPSQLHRSS-LVKNLQNIYFLYEGDPVTHENVKSVDQLLSHHLIYNVSG---PNYD 56
              | || |  ||:||     : |: :||:   |:   |||||: |:| ||||:|     |||
Db         30 QPDPMPDDLHKSSEFTGTMGNMKYLYDDHYVSATKVKSVDKFLAHDLIYNISDKKLKNYD 89

Qy         57 KLKTELKNQEMATLFKDKNVDIYGVEYYHLCYLC--ENAER----SACIYGGVTNHEGNH 110
              |:|||| |:::|  :||: ||:||   ||  ||    :|  :       |:|||:| |||||
Db         90 KVKTELLNEDLAKKYKDEVVDVYGSNYYVNCYFSSKDNVGKVTGGKTCMYGGITKHEGNH 149

Qy        111 LEIP--KKIVVKVSIDGIQSLSFDIETNKKMVTAQELDYKVRKYLTDNKQLYTNGPSKYE 168
               :    : ::|:|  :   ::||:::|:|| |||||| | | :| : | ||    | ||
Db        150 FDNGNLQNVLVRVYENKRNTISFEVQTDKKSVTAQELDIKARNFLINKKNLYEFNSSPYE 209

Qy        169 TGYIKFIPKNKESFWFDFFPEP--EFTQSKYLMIYKDNETLDSNTSQIEVYLTTK 221
              ||||||| |  :||:|  | |  :| ||||||:| ||:|:|| : :|||:||||
Db        210 TGYIKFIENNGNTFWYDMMPAPGDKFDQSKYLMMYNDNKTVDSKSVKIEVHLTTK 264


RESULT 4
ETC2_STAAU
ID   ETC2_STAAU     STANDARD;      PRT;   266 AA.
AC   P34071;
DT   01-FEB-1994 (Rel. 28, Created)
DT   01-FEB-1994 (Rel. 28, Last sequence update)
DT   15-MAR-2004 (Rel. 43, Last annotation update)
DE   Enterotoxin type C-2 precursor (SEC2).
GN   ENTC2.
OS   Staphylococcus aureus.
OC   Bacteria; Firmicutes; Bacillales; Staphylococcus.
OX   NCBI_TaxID=1280;
RN   [1]
RP   SEQUENCE FROM N.A., AND SEQUENCE OF 28-66.
RX   MEDLINE=89277549; PubMed=2543637;
RA   Bohach G.A., Schlievert P.M.;
RT   "Conservation of the biologically active portions of staphylococcal
RT   enterotoxins C1 and C2.";
RL   Infect. Immun. 57:2249-2252(1989).
RN   [2]
RP   X-RAY CRYSTALLOGRAPHY (2.0 ANGSTROMS).
RX   MEDLINE=96027099; PubMed=7582894;
RA   Papageorgiou A.C., Acharya K.R., Shapiro R., Passalacqua E.F.,
RA   Brehm R.D., Tranter H.S.;
RT   "Crystal structure of the superantigen enterotoxin C2 from
RT   Staphylococcus aureus reveals a zinc-binding site.";
RL   Structure 3:769-779(1995).
RN   [3]
RP   X-RAY CRYSTALLOGRAPHY (2.7 ANGSTROMS).
RX   MEDLINE=96022987; PubMed=7552730;
RA   Swaminathan S., Furey W.F. Jr., Pletcher J., Sax M.;
RT   "Residues defining V beta specificity in staphylococcal
RT   enterotoxins.";
RL   Nat. Struct. Biol. 2:680-686(1995).
RN   [4]
RP   COMPARISON OF STRUCTURE OF SEA AND SEC2.
RX   MEDLINE=97334373; PubMed=9191070;
RA   Schad E.M., Papageorgiou A.C., Svensson L.A., Acharya K.R.;
RT   "A structural and functional comparison of staphylococcal
RT   enterotoxins A and C2 reveals remarkable similarity and
RT   dissimilarity.";
RL   J. Mol. Biol. 269:270-280(1997).
CC   -!- FUNCTION: Staphylococcal enterotoxins cause the intoxication
CC       staphylococcal food poisoning syndrome. The illness characterized
CC       by high fever, hypotension, diarrhea, shock, and in some cases
CC       death.
CC   -!- COFACTOR: Binds 1 zinc ion per subunit. The zinc ion is necessary
CC       for the toxin interaction with MHC class II.
CC   -!- SUBCELLULAR LOCATION: Secreted.
CC   -!- SIMILARITY: Belongs to the staphylococcal/streptococcal toxin
CC       family.
DR   PIR; A60114; A60114.
DR   PDB; 1STE; 23-DEC-96.
DR   PDB; 1SE2; 08-MAR-96.
DR   PDB; 1CQV; 19-SEP-01.
DR   PDB; 1I4P; 19-SEP-01.
```

```
DR   PDB; 1I4Q; 19-SEP-01.
DR   PDB; 1I4R; 19-SEP-01.
DR   PDB; 1I4X; 19-SEP-01.
DR   InterPro; IPR008992; Bact_endotox.
DR   InterPro; IPR006177; Bctrl_tox.
DR   InterPro; IPR006123; Stap/Strep_toxin.
DR   InterPro; IPR006126; Staph/Strept_tox.
DR   InterPro; IPR006173; Staph_tox_OB.
DR   Pfam; PF02876; Stap_Strp_tox_C; 1.
DR   Pfam; PF01123; Stap_Strp_toxin; 1.
DR   PRINTS; PR00279; BACTRLTOXIN.
DR   PROSITE; PS00277; STAPH_STREP_TOXIN_1; 1.
DR   PROSITE; PS00278; STAPH_STREP_TOXIN_2; 1.
KW   Enterotoxin; Toxin; Signal; Superantigen; Metal-binding; Zinc;
KW   3D-structure.
FT   SIGNAL        1     27
FT   CHAIN        28    266       ENTEROTOXIN TYPE C-2.
FT   DISULFID    120    137
FT   METAL        36     36       ZINC.
FT   METAL       110    110       ZINC.
FT   METAL       145    145       ZINC.
FT   METAL       149    149       ZINC.
FT   HELIX        35     37
FT   HELIX        41     43
FT   STRAND       44     44
FT   TURN         48     48
FT   HELIX        49     55
FT   STRAND       60     65
FT   STRAND       69     69
FT   TURN         73     74
FT   STRAND       75     79
FT   TURN         83     86
FT   STRAND       90     94
FT   HELIX        98    104
FT   TURN        105    106
FT   STRAND      109    113
FT   STRAND      116    116
FT   TURN        118    119
FT   TURN        128    129
FT   STRAND      136    139
FT   STRAND      142    144
FT   TURN        146    147
FT   STRAND      149    149
FT   HELIX       151    153
FT   STRAND      156    164
FT   TURN        165    166
FT   STRAND      167    176
FT   STRAND      178    178
FT   STRAND      180    182
FT   HELIX       183    198
FT   STRAND      208    216
FT   TURN        218    219
FT   STRAND      222    226
FT   STRAND      232    232
FT   HELIX       237    241
FT   HELIX       242    246
FT   STRAND      249    251
FT   TURN        252    254
FT   STRAND      256    263
SQ   SEQUENCE   266 AA;  30604 MW;  8407FB18536FAC08 CRC64;

  Query Match             46.1%;  Score 543;  DB 1;  Length 266;
  Best Local Similarity   46.4%;  Pred. No. 1.5e-34;
  Matches  109;  Conservative   45;  Mismatches   67;  Indels   14;  Gaps    6;

Qy          1 QQDPDPSQLHRSS-LVKNLQNIYFLYEGDPVTHENVKSVDQLLSHHLIYNVSG---PNYD 56
              | || | :||:||     : |: :||:   |:   | |||: |:| ||||:|     ||| 
Db         30 QPDPTPDELHKSSEFTGTMGNMKYLYDDHYVSATKVMSVDKFLAHDLIYNISDKKLKNYD 89

Qy         57 KLKTELKNQEMATLFKDKNVDIYGVEYYHLCYLC--ENAER----SACIYGGVTNHEGNH 110
              |:|||| |:::|  :||: ||:|| ||  ||    :|  :      |:|||:| ||||| 
Db         90 KVKTELLNEDLAKKYKDEVVDVYGSNYYVNCYFSSKDNVGKVTGGKTCMYGGITKHEGNH 149

Qy        111 LEIP--KKIVVKVSIDGIQSLSFDIETNKKMVTAQELDYKVRKYLTDNKQLYTNGPSKYE 168
              :    : ::::|  :   ::||:::|:|| ||||||| | | :| : | ||    | ||
Db        150 FDNGNLQNVLIRVYENKRNTISFEVQTDKKSVTAQELDIKARNFLINKKNLYEFNSSPYE 209

Qy        169 TGYIKFIPKNKESFWFDFFPEP--EFTQSKYLMIYKDNETLDSNTSQIEVYLTTK 221
              ||||||| |  :||:|  | |  :| ||||||:| ||:|:|| : :|||:|||| 
Db        210 TGYIKFIENNGNTFWYDMMPAPGDKFDQSKYLMMYNDNKTVDSKSVKIEVHLTTK 264


RESULT 5
ETC1_STAAU
ID   ETC1_STAAU     STANDARD;      PRT;   266 AA.
AC   P01553;
DT   21-JUL-1986 (Rel. 01, Created)
DT   01-JAN-1990 (Rel. 13, Last sequence update)
DT   15-MAR-2004 (Rel. 43, Last annotation update)
DE   Enterotoxin type C-1 precursor (SEC1).
GN   ENTC1.
OS   Staphylococcus aureus.
OC   Bacteria; Firmicutes; Bacillales; Staphylococcus.
OX   NCBI_TaxID=1280;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=88038352; PubMed=2823067;
RA   Bohach G.A., Schlievert P.M.;
RT   "Nucleotide sequence of the staphylococcal enterotoxin C1 gene and
RT   relatedness to other pyrogenic toxins.";
RL   Mol. Gen. Genet. 209:15-20(1987).
RN   [2]
RP   SEQUENCE OF 28-266.
RX   MEDLINE=83213327; PubMed=6189824;
RA   Schmidt J.J., Spero L.;
RT   "The complete amino acid sequence of staphylococcal enterotoxin C1.";
RL   J. Biol. Chem. 258:6300-6306(1983).
CC   -!- FUNCTION: Staphylococcal enterotoxins cause the intoxication
CC       staphylococcal food poisoning syndrome. The illness characterized
CC       by high fever, hypotension, diarrhea, shock, and in some cases
CC       death.
CC   -!- SUBCELLULAR LOCATION: Secreted.
CC   -!- SIMILARITY: Belongs to the staphylococcal/streptococcal toxin
CC       family.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; X05815; CAA29260.1; -.
DR   PIR; S06356; ENSAC1.
DR   HSSP; P34071; 1SE2.
DR   InterPro; IPR008992; Bact_endotox.
DR   InterPro; IPR006177; Bctrl_tox.
DR   InterPro; IPR006123; Stap/Strep_toxin.
DR   InterPro; IPR006126; Staph/Strept_tox.
DR   InterPro; IPR006173; Staph_tox_OB.
DR   Pfam; PF02876; Stap_Strp_tox_C; 1.
DR   Pfam; PF01123; Stap_Strp_toxin; 1.
DR   PRINTS; PR00279; BACTRLTOXIN.
DR   PROSITE; PS00277; STAPH_STREP_TOXIN_1; 1.
DR   PROSITE; PS00278; STAPH_STREP_TOXIN_2; 1.
KW   Enterotoxin; Toxin; Signal; Superantigen.
FT   SIGNAL        1     27
FT   CHAIN        28    266       ENTEROTOXIN TYPE C-1.
FT   DISULFID    120    137
FT   CONFLICT    177    177       D -> N (IN REF. 2).
SQ   SEQUENCE   266 AA;  30546 MW;  3A7AB59A8986853B CRC64;

  Query Match             46.0%;  Score 542;  DB 1;  Length 266;
  Best Local Similarity   46.8%;  Pred. No. 1.7e-34;
  Matches  110;  Conservative   44;  Mismatches   67;  Indels   14;  Gaps    6;
```

```
Qy          1 QQDPDPSQLHRSSLVKNL-QNIYFLYEGDPVTHENVKSVDQLLSHHLIYNVSG---PNYD 56
              |  ||  |  :||::|     | :|:  ||:    |:   |||||: |:| ||||:|      |||
Db         30 QPDPTPDELHKASKFTGLMENMKVLYDDHYVSATKVKSVDKFLAHDLIYNISDKKLKNYD 89

Qy         57 KLKTELKNQEMATLFKDKNVDIYGVEYYHLCYLC--ENAER----SACIYGGVTNHEGNH 110
              |:|||| |: :|  :||: ||:||  ||  ||    :|  :       |:|||:| |||||
Db         90 KVKTELLNEGLAKKYKDEVVDVYGSNYYVNCYFSSKDNVGKVTGGKTCMYGGITKHEGNH 149

Qy        111 LEIP--KKIVVKVSIDGIQSLSFDIETNKKMVTAQELDYKVRKYLTDNKQLYTNGPSKYE 168
                  :    : ::::|  :   ::||:::|:|| |||||||  | | :| : | ||    | ||
Db        150 FDNGNLQNVLIRVYENKRNTISFEVQTDKKSVTAQELDIKARNFLINKKNLYEFNSSPYE 209

Qy        169 TGYIKFIPKNKESFWFDFFPEP--EFTQSKYLMIYKDNETLDSNTSQIEVYLTTK 221
              |||||||  |  :||:|  | |  :| ||||||:| ||:|:|| : :|||:||||
Db        210 TGYIKFIENNGNTFWYDMMPAPGDKFDQSKYLMMYNDNKTVDSKSVKIEVHLTTK 264
```

RESULT 6
ETXG_STAAM
ID ETXG_STAAM STANDARD; PRT; 258 AA.
AC O85382;
DT 28-FEB-2003 (Rel. 41, Created)
DT 28-FEB-2003 (Rel. 41, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Enterotoxin type G precursor (SEG).
GN ENTG OR SEG OR SAV1824 OR SA1642.
OS Staphylococcus aureus (strain Mu50 / ATCC 700699),
OS Staphylococcus aureus (strain N315), and
OS Staphylococcus aureus.
OC Bacteria; Firmicutes; Bacillales; Staphylococcus.
OX NCBI_TaxID=158878, 158879, 1280;
RN [1]
RP SEQUENCE FROM N.A., AND PARTIAL SEQUENCE.
RC STRAIN=FRI572;
RX MEDLINE=98298056; PubMed=9632603;
RA Munson S.H., Tremaine M.T., Betley M.J., Welch R.A.;
RT "Identification and characterization of staphylococcal enterotoxin
RT types G and I from Staphylococcus aureus.";
RL Infect. Immun. 66:3337-3348(1998).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=Mu50 / ATCC 700699, and N315;
RX MEDLINE=21311952; PubMed=11418146;
RA Kuroda M., Ohta T., Uchiyama I., Baba T., Yuzawa H., Kobayashi I.,
RA Cui L., Oguchi A., Aoki K.-I., Nagai Y., Lian J.-Q., Ito T.,
RA Kanamori M., Matsumaru H., Maruyama A., Murakami H., Hosoyama A.,
RA Mizutani-Ui Y., Takahashi N.K., Sawano T., Inoue R.-I., Kaito C.,
RA Sekimizu K., Hirakawa H., Kuhara S., Goto S., Yabuzaki J.,
RA Kanehisa M., Yamashita A., Oshima K., Furuya K., Yoshino C., Shiba T.,
RA Hattori M., Ogasawara N., Hayashi H., Hiramatsu K.;
RT "Whole genome sequencing of meticillin-resistant Staphylococcus
RT aureus.";
RL Lancet 357:1225-1240(2001).
CC -!- FUNCTION: Staphylococcal enterotoxins cause the intoxication
CC staphylococcal food poisoning syndrome. The illness characterized
CC by high fever, hypotension, diarrhea, shock, and in some cases
CC death.
CC -!- SUBCELLULAR LOCATION: Secreted.
CC -!- SIMILARITY: Belongs to the staphylococcal/streptococcal toxin
CC family.
CC ---
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC ---
DR EMBL; AF064773; AAC26660.1; -.
DR EMBL; AP003363; BAB57986.1; -.
DR EMBL; AP003135; BAB42910.1; -.
DR PIR; G89968; G89968.
DR HSSP; P01552; 1SBB.
DR InterPro; IPR008992; Bact_endotox.
DR InterPro; IPR006177; Bctrl_tox.
DR InterPro; IPR006123; Stap/Strep_toxin.
DR InterPro; IPR006126; Staph/Strept_tox.
DR InterPro; IPR006173; Staph_tox_OB.
DR Pfam; PF02876; Stap_Strp_tox_C; 1.
DR Pfam; PF01123; Stap_Strp_toxin; 1.
DR PRINTS; PR00279; BACTRLTOXIN.
DR PROSITE; PS00277; STAPH_STREP_TOXIN_1; FALSE_NEG.
DR PROSITE; PS00278; STAPH_STREP_TOXIN_2; 1.
KW Enterotoxin; Toxin; Signal; Superantigen; Complete proteome.
FT SIGNAL 1 25
FT CHAIN 26 258 ENTEROTOXIN TYPE G.
FT DISULFID 116 133 BY SIMILARITY.
SQ SEQUENCE 258 AA; 29940 MW; E2982101701D012C CRC64;

Query Match 39.7%; Score 468; DB 1; Length 258;
Best Local Similarity 43.8%; Pred. No. 7.4e-29;
Matches 102; Conservative 39; Mismatches 78; Indels 14; Gaps 4;

```
Qy          1 QQDPDPSQLHRSSLVKN----LQNIYFLYEGDPVTHENVKSVDQLLSHHLIYNVSGPNYD 56
              |  ||    :|:: |  ||     : |:  ||   ||     | :  | ||| ||: :    :|:
Db         26 QPDPKLDELNKVSDYKNNKGTMGNVMNLYTSPPVEGRGVINSRQFLSHDLIFPIEYKSYN 85

Qy         57 KLKTELKNQEMATLFKDKNVDIYGVEYYHLCYLCENAE------RSACIYGGVTNHEGNH 110
              ::||||:| |:|  :||| |||:|| |:: | : ::           |:|||:| :   :
Db         86 EVKTELENTELANNYKDKKVDIFGVPYFYTCIIPKSEPDINQNFGGCCMYGGLTFNSSEN 145

Qy        111 LEIPKKIVVKVSIDGIQSLSFDIETNKKMVTAQELDYKVRKYLTDNKQLYTNGPSKYETG 170
              |  | | |:|:||  ||| | | ||| ||| |||||| | :||  |:||     | :|:|
Db        146 -ERDKLITVQVTIDNRQSLGFTITTNKNMVTIQELDYKARHWLTKEKKLYEFDGSAFESG 204

Qy        171 YIKFIPKNKESFWFDFFPEPE---FTQSKYLMIYKDNETLDSNTSQIEVYLTT 220
              ||||  ||  ||||| ||: |   |   |:| || ||: :|| : ::||:| |
Db        205 YIKFTEKNNTSFWFDLFPKKELVPFVPYKFLNIYGDNKVVDSKSIKMEVFLNT 257
```

RESULT 7
ETXD_STAAU
ID ETXD_STAAU STANDARD; PRT; 258 AA.
AC P20723;
DT 01-FEB-1991 (Rel. 17, Created)
DT 01-FEB-1991 (Rel. 17, Last sequence update)
DT 15-MAR-2004 (Rel. 43, Last annotation update)
DE Enterotoxin type D precursor (SED).
GN ENTD.
OS Staphylococcus aureus.
OC Bacteria; Firmicutes; Bacillales; Staphylococcus.
OX NCBI_TaxID=1280;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=89359112; PubMed=2549000;
RA Bayles K.W., Iandolo J.J.;
RT "Genetic and molecular analyses of the gene encoding staphylococcal
RT enterotoxin D.";
RL J. Bacteriol. 171:4799-4806(1989).
RN [2]
RP SEQUENCE FROM N.A., AND X-RAY CRYSTALLOGRAPHY (2.3 ANGSTROMS).
RC STRAIN=ATCC 23235;
RX MEDLINE=97157473; PubMed=9003758;
RA Sundstroem M., Abrahmsen L., Antonsson P., Mehindate K., Mourad W.,
RA Dohlsten M.;
RT "The crystal structure of staphylococcal enterotoxin type D reveals
RT Zn2+-mediated homodimerization.";
RL EMBO J. 15:6832-6840(1996).
CC -!- FUNCTION: Staphylococcal enterotoxins cause the intoxication
CC staphylococcal food poisoning syndrome. The illness characterized
CC by high fever, hypotension, diarrhea, shock, and in some cases
CC death.
CC -!- SUBUNIT: Homodimer; zinc-dependent.
CC -!- SUBCELLULAR LOCATION: Secreted.

```
CC   -!- SIMILARITY: Belongs to the staphylococcal/streptococcal toxin
CC       family.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; M28521; AAB06195.1; -.
DR   PIR; A33953; A33953.
DR   HSSP; P13163; 1SXT.
DR   InterPro; IPR008992; Bact_endotox.
DR   InterPro; IPR006177; Bctrl_tox.
DR   InterPro; IPR006123; Stap/Strep_toxin.
DR   InterPro; IPR006126; Staph/Strept_tox.
DR   InterPro; IPR006173; Staph_tox_OB.
DR   Pfam; PF02876; Stap_Strp_tox_C; 1.
DR   Pfam; PF01123; Stap_Strp_toxin; 1.
DR   PRINTS; PR00279; BACTRLTOXIN.
DR   PROSITE; PS00277; STAPH_STREP_TOXIN_1; 1.
DR   PROSITE; PS00278; STAPH_STREP_TOXIN_2; 1.
KW   Enterotoxin; Toxin; Signal; Superantigen; Zinc.
FT   SIGNAL        1     25
FT   CHAIN        26    258       ENTEROTOXIN TYPE D.
FT   METAL       212    212       ZINC.
FT   METAL       250    250       ZINC.
FT   METAL       252    252       ZINC.
FT   VARIANT     114    114       P -> A (IN STRAIN ATCC 23235).
SQ   SEQUENCE   258 AA;  29746 MW;  4F7C6A28D42597FD CRC64;

  Query Match             28.3%;  Score 333.5;  DB 1;  Length 258;
  Best Local Similarity   36.7%;  Pred. No. 1.4e-18;
  Matches   83;  Conservative   42;  Mismatches   86;  Indels   15;  Gaps    7;

Qy          8 QLHR-----SSLVKNLQNIYFLYEGDPVTHENVKSVDQLLSHHLIYN---VSGPNYDKLK 59
              :||:     |: : |::: |   : :|:  ||  : || | : |:|       |:: |
Db         36 ELHKKSELSSTALNNMKHSY--ADKNPIIGENKSTGDQFLENTLLYKKFFTDLINFEDLL 93

Qy         60 TELKNQEMATLFKDKNVDIYGVEYYHLCYLCENAERSACIYGGVTNHEGNHLEIPKKIVV 119
                ::|||  || ||||:| : |   ||  |  :|:|| ||||| |||| |:  ||| :
Db         94 INFNSKEMAQHFKSKNVDVYPIRYSINCYGGE-IDRTACTYGGVTPHEGNKLKERKKIPI 152

Qy        120 KVSIDGIQ-SLSFD-IETNKKMVTAQELDYKVRKYLTDNKQLYTNGP--SKYETGYIKFI 175
               : |:|:|  :| | ::|:|| || |||| : |:||  : :|| |      | : | |:|
Db        153 NLWINGVQKEVSLDKVQTDKKNVTVQELDAQARRYLQKDLKLYNNDTLGGKIQRGKIEFD 212

Qy        176 PKNKESFWFDFFPEPEFTQSKYLMIYKDNETLDSNTSQIEVYLTTK 221
                 :     :| |          | | || ||:|| :    |::||  |
Db        213 SSDGSKVSYDLFDVKGDFPEKQLRIYSDNKTLSTEHLHIDIYLYEK 258


RESULT 8
ETXA_STAAW
ID   ETXA_STAAW     STANDARD;      PRT;   257 AA.
AC   P13163;
DT   01-JAN-1990 (Rel. 13, Created)
DT   01-JAN-1990 (Rel. 13, Last sequence update)
DT   15-MAR-2004 (Rel. 43, Last annotation update)
DE   Enterotoxin type A precursor (SEA).
GN   ENTA OR MW1889.
OS   Staphylococcus aureus (strain MW2), and
OS   Staphylococcus aureus.
OC   Bacteria; Firmicutes; Bacillales; Staphylococcus.
OX   NCBI_TaxID=196620, 1280;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=MW2;
RX   MEDLINE=22040717; PubMed=12044378;
RA   Baba T., Takeuchi F., Kuroda M., Yuzawa H., Aoki K.-I., Oguchi A.,
RA   Nagai Y., Iwama N., Asano K., Naimi T., Kuroda H., Cui L.,
RA   Yamamoto K., Hiramatsu K.;
RT   "Genome and virulence determinants of high virulence community-
RT   acquired MRSA.";
RL   Lancet 359:1819-1827(2002).
RN   [2]
RP   SEQUENCE FROM N.A.
RC   STRAIN=FRI337;
RX   MEDLINE=88086892; PubMed=3335483;
RA   Betley M.J., Mekalanos J.J.;
RT   "Nucleotide sequence of the type A staphylococcal enterotoxin gene.";
RL   J. Bacteriol. 170:34-41(1988).
RN   [3]
RP   SEQUENCE OF 25-257.
RX   MEDLINE=87222293; PubMed=3584106;
RA   Huang I.-Y., Hughes J.L., Bergdoll M.S., Schantz E.J.;
RT   "Complete amino acid sequence of staphylococcal enterotoxin A.";
RL   J. Biol. Chem. 262:7006-7013(1987).
RN   [4]
RP   X-RAY CRYSTALLOGRAPHY (1.9 ANGSTROMS).
RX   MEDLINE=95354648; PubMed=7628431;
RA   Schad E.M., Zaitseva I., Zaitsev V.N., Dohlsten M., Kalland T.,
RA   Schlievert P.M., Ohlendorf D.H., Svensson L.A.;
RT   "Crystal structure of the superantigen staphylococcal enterotoxin
RT   type A.";
RL   EMBO J. 14:3292-3301(1995).
RN   [5]
RP   X-RAY CRYSTALLOGRAPHY (2.7 ANGSTROMS).
RX   MEDLINE=97113025; PubMed=8943278;
RA   Sundstroem M., Hallen D., Svensson A., Schad E., Dohlsten M.,
RA   Abrahmsen L.;
RT   "The Co-crystal structure of staphylococcal enterotoxin type A with
RT   Zn2+ at 2.7-A resolution. Implications for major histocompatibility
RT   complex class II binding.";
RL   J. Biol. Chem. 271:32212-32216(1996).
RN   [6]
RP   3D-STRUCTURE MODELING.
RX   MEDLINE=96022987; PubMed=7552730;
RA   Swaminathan S., Furey W.F. Jr., Pletcher J., Sax M.;
RT   "Residues defining V beta specificity in staphylococcal
RT   enterotoxins.";
RL   Nat. Struct. Biol. 2:680-686(1995).
RN   [7]
RP   COMPARISON OF STRUCTURE OF SEA AND SEC2.
RX   MEDLINE=97334373; PubMed=9191070;
RA   Schad E.M., Papageorgiou A.C., Svensson L.A., Acharya K.R.;
RT   "A structural and functional comparison of staphylococcal
RT   enterotoxins A and C2 reveals remarkable similarity and
RT   dissimilarity.";
RL   J. Mol. Biol. 269:270-280(1997).
CC   -!- FUNCTION: Staphylococcal enterotoxins cause the intoxication
CC       staphylococcal food poisoning syndrome. The illness characterized
CC       by high fever, hypotension, diarrhea, shock, and in some cases
CC       death.
CC   -!- COFACTOR: Binds 1 zinc ion per subunit. The zinc ion is necessary
CC       for the toxin interaction with MHC class II.
CC   -!- SUBUNIT: Monomer.
CC   -!- SUBCELLULAR LOCATION: Secreted.
CC   -!- MISCELLANEOUS: THIS TOXIN SEEMS TO BE CODED BY A BACTERIOPHAGE.
CC   -!- SIMILARITY: Belongs to the staphylococcal/streptococcal toxin
CC       family.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; AP004828; BAB95754.1; -.
DR   EMBL; M18970; AAA26681.1; -.
DR   PIR; A28664; A28664.
DR   PDB; 1ESF; 11-JUL-96.
```

```
DR   PDB; 1SXT; 19-NOV-97.
DR   PDB; 1DYQ; 21-FEB-02.
DR   PDB; 1I4G; 21-MAR-01.
DR   PDB; 1I4H; 21-MAR-01.
DR   PDB; 1LO5; 18-DEC-02.
DR   PDB; 1SEA; 15-OCT-95.
DR   InterPro; IPR008992; Bact_endotox.
DR   InterPro; IPR006177; Bctrl_tox.
DR   InterPro; IPR006123; Stap/Strept_toxin.
DR   InterPro; IPR006126; Staph/Strept_tox.
DR   InterPro; IPR006173; Staph_tox_OB.
DR   Pfam; PF02876; Stap_Strp_tox_C; 1.
DR   Pfam; PF01123; Stap_Strp_toxin; 1.
DR   PRINTS; PR00279; BACTRLTOXIN.
DR   PROSITE; PS00277; STAPH_STREP_TOXIN_1; 1.
DR   PROSITE; PS00278; STAPH_STREP_TOXIN_2; 1.
KW   Enterotoxin; Toxin; Signal; Superantigen; Metal-binding; Zinc;
KW   3D-structure; Complete proteome.
FT   SIGNAL        1    24
FT   CHAIN        25   257       ENTEROTOXIN TYPE A.
FT   DISULFID    120   130
FT   METAL       211   211       ZINC.
FT   METAL       249   249       ZINC.
FT   METAL       251   251       ZINC.
FT   CONFLICT    242   242       T -> S (IN REF. 3).
FT   HELIX        28    31
FT   TURN         32    33
FT   HELIX        39    41
FT   TURN         44    45
FT   HELIX        46    55
FT   TURN         56    56
FT   STRAND       59    65
FT   STRAND       69    69
FT   TURN         73    74
FT   STRAND       75    78
FT   TURN         79    80
FT   STRAND       90    94
FT   HELIX        98   104
FT   TURN        105   106
FT   STRAND      108   116
FT   TURN        118   119
FT   TURN        125   126
FT   STRAND      128   132
FT   STRAND      135   137
FT   TURN        139   140
FT   STRAND      142   149
FT   STRAND      151   155
FT   TURN        156   157
FT   STRAND      158   160
FT   HELIX       164   166
FT   STRAND      167   171
FT   STRAND      173   175
FT   HELIX       176   191
FT   TURN        193   194
FT   HELIX       197   199
FT   TURN        200   200
FT   STRAND      205   212
FT   STRAND      218   221
FT   TURN        222   223
FT   HELIX       230   233
FT   HELIX       234   237
FT   TURN        238   239
FT   STRAND      242   244
FT   STRAND      248   255
SQ   SEQUENCE   257 AA;  29669 MW;  ADEBF5BCA1F14677 CRC64;

  Query Match             24.9%;  Score 293.5;  DB 1;  Length 257;
  Best Local Similarity   33.0%;  Pred. No. 1.5e-15;
  Matches   73;  Conservative   43;  Mismatches   94;  Indels   11;  Gaps    6;

Qy          7 SQLHRSSLVKNLQNIYFLYEGDPVTHENVKSVDQLLSHHLIYN---VSGPNYDKLKTELK 63
              |:|  ::|  ||: ||  |  :    || :| || | | :::         |: |  :
Db         40 SELQGTAL-GNLKQIY--YYNEKAKTENKESHDQFLQHTILFKGFFTDHSWYNDLLVDFD 96

Qy         64 NQEMATLFKDKNVDIYGVEYYHLCYLCENAERSACIYGGVTNHEGNHLEIPKKIVVKVSI 123
              ::::   :| | ||:|| | : |       ::||:||||| |: | |   ||: : : :
Db         97 SKDIVDKYKGKKVDLYGAYYGYQC-AGGTPNKTACMYGGVTLHDNNRLTEEKKVPINLWL 155

Qy        124 DGIQSL--SFDIETNKKMVTAQELDYKVRKYLTDNKQLYTNG--PSKYETGYIKFIPKNK 179
              || |:      ::|||| || |||| : |:|| :   || :     | : | | |    :
Db        156 DGKQNTVPLETVKTNKKNVTVQELDLQARRYLQEKYNLYNSDVFDGKVQRGLIVFHTSTE 215

Qy        180 ESFWFDFFPEPEFTQSKYLMIYKDNETLDSNTSQIEVYLTT 220
              |  :| |        :  | ||:||:|::|    |::|| |
Db        216 PSVNYDLFGAQGQYSNTLLRIYRDNKTINSENMHIDIYLYT 256


RESULT 9
ETXE_STAAU
ID   ETXE_STAAU     STANDARD;      PRT;   257 AA.
AC   P12993;
DT   01-JAN-1990 (Rel. 13, Created)
DT   01-JAN-1990 (Rel. 13, Last sequence update)
DT   15-MAR-2004 (Rel. 43, Last annotation update)
DE   Enterotoxin type E precursor (SEE).
GN   ENTE.
OS   Staphylococcus aureus.
OC   Bacteria; Firmicutes; Bacillales; Staphylococcus.
OX   NCBI_TaxID=1280;
RN   [1]
RP   SEQUENCE FROM N.A., AND SEQUENCE OF 28-74.
RC   STRAIN=MJB265;
RX   MEDLINE=88257005; PubMed=3384800;
RA   Couch J.L., Soltis M.T., Betley M.J.;
RT   "Cloning and nucleotide sequence of the type E staphylococcal
RT   enterotoxin gene.";
RL   J. Bacteriol. 170:2954-2960(1988).
RN   [2]
RP   3D-STRUCTURE MODELING.
RX   MEDLINE=96022987; PubMed=7552730;
RA   Swaminathan S., Furey W.F. Jr., Pletcher J., Sax M.;
RT   "Residues defining V beta specificity in staphylococcal
RT   enterotoxins.";
RL   Nat. Struct. Biol. 2:680-686(1995).
CC   -!- FUNCTION: Staphylococcal enterotoxins cause the intoxication
CC       staphylococcal food poisoning syndrome. The illness characterized
CC       by high fever, hypotension, diarrhea, shock, and in some cases
CC       death.
CC   -!- COFACTOR: Binds 1 zinc ion per subunit. The zinc ion is necessary
CC       for the toxin interaction with MHC class II (By similarity).
CC   -!- SUBCELLULAR LOCATION: Secreted.
CC   -!- SIMILARITY: Belongs to the staphylococcal/streptococcal toxin
CC       family.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; M21319; AAA26617.1; -.
DR   PIR; A28179; A28179.
DR   PDB; 1SEE; 15-OCT-95.
DR   InterPro; IPR008992; Bact_endotox.
DR   InterPro; IPR006177; Bctrl_tox.
DR   InterPro; IPR006123; Stap/Strep_toxin.
DR   InterPro; IPR006126; Staph/Strept_tox.
DR   InterPro; IPR006173; Staph_tox_OB.
DR   Pfam; PF02876; Stap_Strp_tox_C; 1.
DR   Pfam; PF01123; Stap_Strp_toxin; 1.
DR   PRINTS; PR00279; BACTRLTOXIN.
DR   PROSITE; PS00277; STAPH_STREP_TOXIN_1; 1.
DR   PROSITE; PS00278; STAPH_STREP_TOXIN_2; 1.
KW   Enterotoxin; Toxin; Signal; Superantigen; Metal-binding; Zinc;
```

```
KW   3D-structure.
FT   SIGNAL       1    27
FT   CHAIN       28   257       ENTEROTOXIN TYPE E.
FT   METAL      211   211       ZINC (BY SIMILARITY).
FT   METAL      249   249       ZINC (BY SIMILARITY).
FT   METAL      251   251       ZINC (BY SIMILARITY).
FT   HELIX       33    35
FT   HELIX       39    41
FT   STRAND      42    42
FT   TURN        46    47
FT   HELIX       48    51
FT   HELIX       52    54
FT   STRAND      59    64
FT   TURN        66    67
FT   STRAND      69    69
FT   TURN        73    74
FT   STRAND      75    79
FT   TURN        83    84
FT   STRAND      90    94
FT   HELIX       98   104
FT   TURN       105   106
FT   STRAND     109   113
FT   STRAND     115   116
FT   TURN       118   119
FT   STRAND     128   132
FT   STRAND     135   137
FT   TURN       139   140
FT   STRAND     142   142
FT   STRAND     152   152
FT   STRAND     154   155
FT   TURN       156   157
FT   STRAND     163   163
FT   STRAND     171   171
FT   STRAND     173   175
FT   HELIX      176   190
FT   TURN       191   191
FT   STRAND     203   211
FT   STRAND     219   221
FT   STRAND     227   227
FT   HELIX      234   239
FT   STRAND     242   244
FT   TURN       245   247
FT   STRAND     249   257
SQ   SEQUENCE   257 AA;  29358 MW;  27EDA94B97770CE3 CRC64;

  Query Match           24.7%;  Score 291.5;  DB 1;  Length 257;
  Best Local Similarity   35.1%;  Pred. No. 2.2e-15;
  Matches   79;  Conservative   42;  Mismatches   85;  Indels   19;  Gaps   10;

Qy          7 SQLHRSSLVKNLQNIYFLYEGDPVTHENVKSVDQLLSHHLIYN---VSGPNYDKLKTELK 63
              |:| |::|  ||: ||: |    :| || :| || | : |::       | |: |  :|
Db         40 SELQRNAL-SNLRQIYY-YNEKAIT-ENKESDDQFLENTLLFKGFFTGHPWYNDLLVDLG 96

Qy         64 NQEMATLFKDKNVDIYGVEYYHLCYLCENAERSACIYGGVTNHEGNHLEIPKKIVVKVSI 123
              :::    :| | ||:||  | : |       ::||:||||| |: | |   ||: : : |
Db         97 SKDATNKYKGKKVDLYGAYYGYQC-AGGTPNKTACMYGGVTLHDNNRLTEEKKVPINLWI 155

Qy        124 DGIQ-SLSFD-IETNKKMVTAQELDYKVRKYLTDNKQLYTNGP--SKYETGYIKFIPKNK 179
              || | ::  | ::|:|| || |||| : | ||     || :     | : | | |
Db        156 DGKQTTVPIDKVKTSKKEVTVQELDLQARHYLHGKFGLYNSDSFGGKVQRGLIVFHSSEG 215

Qy        180 ESFWFDFPPEPEFTQSKY----LMIYKDNETLDSNTSQIEVYLTT 220
               :  :| |      | :|    | ||:||:|::|    |::|| |
Db        216 STVSYDLFD----AQGQYPDTLLRIYRDNKTINSENLHIDLYLYT 256


RESULT 10
ETXH_STAAW
ID   ETXH_STAAW     STANDARD;      PRT;   241 AA.
AC   Q53585;
DT   15-MAR-2004 (Rel. 43, Created)
DT   15-MAR-2004 (Rel. 43, Last sequence update)
DT   15-MAR-2004 (Rel. 43, Last annotation update)
DE   Enterotoxin type H precursor (SEH).
GN   ENTH OR SEH OR MW0051.
OS   Staphylococcus aureus (strain MW2), and
OS   Staphylococcus aureus.
OC   Bacteria; Firmicutes; Bacillales; Staphylococcus.
OX   NCBI_TaxID=196620, 1280;
RN   [1]
RP   SEQUENCE FROM N.A., PARTIAL SEQUENCE, AND CHARACTERIZATION.
RC   STRAIN=D4508;
RX   MEDLINE=95053699; PubMed=7964453;
RA   Ren K., Bannan J.D., Pancholi V., Cheung A.L., Robbins J.C.,
RA   Fischetti V.A., Zabriskie J.B.;
RT   "Characterization and biological properties of a new staphylococcal
RT   exotoxin.";
RL   J. Exp. Med. 180:1675-1683(1994).
RN   [2]
RP   SEQUENCE FROM N.A.
RC   STRAIN=MW2;
RX   MEDLINE=22040717; PubMed=12044378;
RA   Baba T., Takeuchi F., Kuroda M., Yuzawa H., Aoki K.-I., Oguchi A.,
RA   Nagai Y., Iwama N., Asano K., Naimi T., Kuroda H., Cui L.,
RA   Yamamoto K., Hiramatsu K.;
RT   "Genome and virulence determinants of high virulence community-
RT   acquired MRSA.";
RL   Lancet 359:1819-1827(2002).
RN   [3]
RP   X-RAY CRYSTALLOGRAPHY (1.69 ANGSTROMS).
RX   MEDLINE=20444256; PubMed=10986116;
RA   Haekansson M., Petersson K., Nilsson H., Forsberg G., Bjoerk P.,
RA   Antonsson P., Svensson L.A.;
RT   "The crystal structure of staphylococcal enterotoxin H: implications
RT   for binding properties to MHC class II and TcR molecules.";
RL   J. Mol. Biol. 302:527-537(2000).
CC   -!- FUNCTION: Staphylococcal enterotoxins cause the intoxication
CC       staphylococcal food poisoning syndrome. The illness characterized
CC       by high fever, hypotension, diarrhea, shock, and in some cases
CC       death.
CC   -!- COFACTOR: Binds 1 zinc ion per subunit. The zinc ion is necessary
CC       for the toxin interaction with MHC class II.
CC   -!- SUBCELLULAR LOCATION: Secreted.
CC   -!- SIMILARITY: Belongs to the staphylococcal/streptococcal toxin
CC       family.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; U11702; AAA19777.1; -.
DR   EMBL; AP004822; BAB93916.1; -.
DR   PDB; 1ENF; 10-JAN-01.
DR   PDB; 1EWC; 10-JAN-01.
DR   PDB; 1F77; 10-JAN-01.
DR   PDB; 1HXY; 27-JUN-01.
DR   InterPro; IPR008992; Bact_endotox.
DR   InterPro; IPR006177; BctrI_tox.
DR   InterPro; IPR006123; Stap/Strep_toxin.
DR   InterPro; IPR006173; Staph_tox_OB.
DR   Pfam; PF01123; Stap_Strp_toxin; 1.
DR   Pfam; PF02876; Stap_Strp_tox_C; 1.
DR   PRINTS; PR00279; BACTRLTOXIN.
DR   PROSITE; PS00277; STAPH_STREP_TOXIN_1; FALSE_NEG.
DR   PROSITE; PS00278; STAPH_STREP_TOXIN_2; 1.
KW   Enterotoxin; Toxin; Signal; Superantigen; Metal-binding; Zinc;
KW   3D-structure; Complete proteome.
FT   SIGNAL       1    24
FT   CHAIN       25   241       ENTEROTOXIN TYPE H.
FT   METAL      230   230       ZINC.
FT   METAL      232   232       ZINC.
FT   DISULFID   106   116
```

```
SQ   SEQUENCE   241 AA;  27858 MW;  70F77985877616CE CRC64;

  Query Match             23.9%;  Score 281.5;  DB 1;  Length 241;
  Best Local Similarity   38.0%;  Pred. No. 1.2e-14;
  Matches   84;  Conservative   33;  Mismatches   87;  Indels   17;  Gaps   10;

Qy          9 LHRSSLVKN--LQNIYFLYEGDPVTHENVKSVDQLLSHHLIYNVSGPNYDKLKTELKNQE 66
              ||  | : :  | | |  |    |   ||:|| :      ||:   | : : |: :    :
Db         27 LHDKSELTDLALANAYGQY-NHPFIKENIKSDEISGEKDLIFRNQGDSGNDLRVKFATAD 85

Qy         67 MATLFKDKNVDIYGVEYYHLC-YLCENAERSACIYGGVT-NHEGNHLEIPKKIVVKVSID 124
              :|  ||:|||||||  :|: |  : ||   | |:||| | | |   |   : |   | :|
Db         86 LAQKFKNKNVDIYGASFYYKCEKISENI--SECLYGGTTLNSE--KLAQERVIGANVWVD 141

Qy        125 GIQSLSFDIETNKKMVTAQELDYKVRKYLTDNKQLYTNGPSKYETGYIKFIPKNKESFWF 184
              |||  :  | |||| || |||| |:|| |:|  ::|    |:   | |:|  |    : |
Db        142 GIQKETELIRTNKKNVTLQELDIKIRKILSDKYKIYYK-DSEISKGLIEFDMKTPRDYSF 200

Qy        185 DFFP---EPEFTQSKYLMIYKDNETLDS-NTSQIEVYLTTK 221
              | :    | ::   |   ||:||:|| | : | |:| | ||
Db        201 DIYDLKGENDYEIDK---IYEDNKTLKSDDISHIDVNLYTK 238


RESULT 11
SPEH_STRPY
ID   SPEH_STRPY     STANDARD;      PRT;   236 AA.
AC   Q9X5C8;
DT   16-OCT-2001 (Rel. 40, Created)
DT   16-OCT-2001 (Rel. 40, Last sequence update)
DT   10-OCT-2003 (Rel. 42, Last annotation update)
DE   Exotoxin type H precursor (SPE H).
GN   SPEH OR SPY1008.
OS   Streptococcus pyogenes.
OC   Bacteria; Firmicutes; Lactobacillales; Streptococcaceae;
OC   Streptococcus.
OX   NCBI_TaxID=1314;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=M15;
RX   MEDLINE=99093428; PubMed=9874566;
RA   Proft T., Moffatt S.L., Berkahn C.J., Fraser J.D.;
RT   "Identification and characterization of novel superantigens from
RT   Streptococcus pyogenes.";
RL   J. Exp. Med. 189:89-102(1999).
RN   [2]
RP   SEQUENCE FROM N.A.
RC   STRAIN=SF370 / ATCC 700294 / Serotype M1;
RX   MEDLINE=21192684; PubMed=11296296;
RA   Ferretti J.J., McShan W.M., Ajdic D.J., Savic D.J., Savic G., Lyon K.,
RA   Primeaux C., Sezate S., Suvorov A.N., Kenton S., Lai H.S., Lin S.P.,
RA   Qian Y., Jia H.G., Najar F.Z., Ren Q., Zhu H., Song L., White J.,
RA   Yuan X., Clifton S.W., Roe B.A., McLaughlin R.;
RT   "Complete genome sequence of an M1 strain of Streptococcus pyogenes.";
RL   Proc. Natl. Acad. Sci. U.S.A. 98:4658-4663(2001).
CC   -!- FUNCTION: Mitogenic for human peripheral blood lymphocytes.
CC   -!- SUBUNIT: Binds to major histocompatibility complex class II beta
CC       chain.
CC   -!- SUBCELLULAR LOCATION: Secreted.
CC   -!- SIMILARITY: Belongs to the staphylococcal/streptococcal toxin
CC       family.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; AF124500; AAD30989.1; -.
DR   EMBL; AE006546; AAK33907.1; -.
DR   PDB; 1ET9; 24-MAY-00.
DR   PDB; 1EU4; 24-MAY-00.
DR   InterPro; IPR008992; Bact_endotox.
DR   InterPro; IPR006123; Stap/Strep_toxin.
DR   InterPro; IPR006126; Staph/Strept_tox.
DR   InterPro; IPR006173; Staph_tox_OB.
DR   Pfam; PF02876; Stap_Strp_tox_C; 1.
DR   Pfam; PF01123; Stap_Strp_toxin; 1.
DR   PROSITE; PS00277; STAPH_STREP_TOXIN_1; FALSE_NEG.
DR   PROSITE; PS00278; STAPH_STREP_TOXIN_2; 1.
KW   Toxin; Signal; Complete proteome; 3D-structure.
FT   SIGNAL        1     32       POTENTIAL.
FT   CHAIN        33    236       EXOTOXIN TYPE H.
SQ   SEQUENCE   236 AA;  27485 MW;  16352923907AD40D CRC64;

  Query Match             17.8%;  Score 209.5;  DB 1;  Length 236;
  Best Local Similarity   31.4%;  Pred. No. 3.5e-09;
  Matches   66;  Conservative   32;  Mismatches   93;  Indels   19;  Gaps    9;

Qy         17 NLQNIYFLYEGDP--VTHENVK-SVDQLLSHHLIYNVSGPNYDKLKTELKNQEMATLFKD 73
              |  |:  ||: |   :  :::| | | : || | |:|   |   |     :   ::  |||
Db         39 NRHNLESLYKHDSNLIEADSIKNSPDIVTSHMLKYSVKDKN---LSVFFEKDWISQEFKD 95

Qy         74 KNVDIYGVEYYHLCYLCENAERSACIYGGV--TNHEGNHLEIPKKIVVKVSIDGIQSLSF 131
              | |||| :    :|  |       |  :||:  || |   :::|     | |      |
Db         96 KEVDIYALSAQEVCE-CPGKRYEA--FGGITLTNSEKKEIKVP----VNVWDKSKQQPPM 148

Qy        132 DIETNKKMVTAQELDYKVRKYLTDNKQLYTNGPSKYETGYIKFIPKNKESFWFD--FFPE 189
              |  ||  |||||:| ||||  |       :| |   ||  | :     : :   ||  :|
Db        149 FITVNKPKVTAQEVDIKVRKLLIKKYDIYNNREQKYSKGTVTLDLNSGKDIVFDLYYFGN 208

Qy        190 PEFTQSKYLMIYKDNETLDSNTSQIEVYLT 219
              :|  :  | || :|| :||    ::| ::
Db        209 GDF--NSMLKIYSNNERIDSTQFHVDVSIS 236


RESULT 12
SPEC_STRPY
ID   SPEC_STRPY     STANDARD;      PRT;   235 AA.
AC   P13380;
DT   01-JAN-1990 (Rel. 13, Created)
DT   16-OCT-2001 (Rel. 40, Last sequence update)
DT   15-MAR-2004 (Rel. 43, Last annotation update)
DE   Exotoxin type C precursor (SPE C).
GN   SPEC OR SPY0711.
OS   Streptococcus pyogenes.
OC   Bacteria; Firmicutes; Lactobacillales; Streptococcaceae;
OC   Streptococcus.
OX   NCBI_TaxID=1314;
RN   [1]
RP   SEQUENCE FROM N.A., AND SEQUENCE OF 28-52.
RC   STRAIN=T18P / MGAS 1585;
RX   MEDLINE=88314303; PubMed=3045005;
RA   Goshorn S.C., Schlievert P.M.;
RT   "Nucleotide sequence of streptococcal pyrogenic exotoxin type C.";
RL   Infect. Immun. 56:2518-2520(1988).
RN   [2]
RP   REVISIONS TO 21-26.
RC   STRAIN=T18P / MGAS 1585;
RX   MEDLINE=92363541; PubMed=1500157;
RA   Kapur V., Nelson K., Schlievert P.M., Selander R.K., Musser J.M.;
RT   "Molecular population genetic evidence of horizontal spread of two
RT   alleles of the pyrogenic exotoxin C gene (speC) among pathogenic
RT   clones of Streptococcus pyogenes.";
RL   Infect. Immun. 60:3513-3517(1992).
RN   [3]
RP   SEQUENCE FROM N.A.
RC   STRAIN=SF370 / ATCC 700294 / Serotype M1;
RX   MEDLINE=21192684; PubMed=11296296;
RA   Ferretti J.J., McShan W.M., Ajdic D.J., Savic D.J., Savic G., Lyon K.,
RA   Primeaux C., Sezate S., Suvorov A.N., Kenton S., Lai H.S., Lin S.P.,
RA   Qian Y., Jia H.G., Najar F.Z., Ren Q., Zhu H., Song L., White J.,
RA   Yuan X., Clifton S.W., Roe B.A., McLaughlin R.;
RT   "Complete genome sequence of an M1 strain of Streptococcus
RT   pyogenes.";
```

```
RL   Proc. Natl. Acad. Sci. U.S.A. 98:4658-4663(2001).
RN   [4]
RP   X-RAY CRYSTALLOGRAPHY (2.4 ANGSTROMS) OF 30-235.
RX   MEDLINE=97397352; PubMed=9253413;
RA   Roussel A., Anderson B.F., Baker H.M., Fraser J.D., Baker E.N.;
RT   "Crystal structure of the streptococcal superantigen SPE-C:
RT   dimerization and zinc binding suggest a novel mode of interaction
RT   with MHC class II molecules.";
RL   Nat. Struct. Biol. 4:635-643(1997).
CC   -!- FUNCTION: Causative agent of the symptoms associated with scarlet
CC       fever, have been associated with streptococcal toxic shocK-like
CC       disease and may play a role in the early events of rheumatic
CC       fever.
CC   -!- SUBUNIT: Binds to major histocompatibility complex class II beta
CC       chain.
CC   -!- MISCELLANEOUS: THIS TOXIN SEEMS TO BE CODED BY A BACTERIOPHAGE.
CC   -!- SIMILARITY: Belongs to the staphylococcal/streptococcal toxin
CC       family.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; M35514; AAA27017.1; ALT_SEQ.
DR   EMBL; M97156; AAB59091.1; -.
DR   EMBL; M97157; AAB59092.1; -.
DR   EMBL; AE006523; AAK33664.1; -.
DR   PIR; A30509; A30509.
DR   PIR; A44799; A44799.
DR   PDB; 1AN8; 29-APR-98.
DR   PDB; 1KTK; 07-JUN-02.
DR   InterPro; IPR008992; Bact_endotox.
DR   InterPro; IPR006177; Bctrl_tox.
DR   InterPro; IPR006123; Stap/Strep_toxin.
DR   InterPro; IPR006126; Staph/Strept_tox.
DR   InterPro; IPR006173; Staph_tox_OB.
DR   Pfam; PF02876; Stap_Strp_tox_C; 1.
DR   Pfam; PF01123; Stap_Strp_toxin; 1.
DR   PRINTS; PR00279; BACTRLTOXIN.
DR   PROSITE; PS00277; STAPH_STREP_TOXIN_1; 1.
DR   PROSITE; PS00278; STAPH_STREP_TOXIN_2; 1.
KW   Toxin; Signal; 3D-structure; Complete proteome.
FT   SIGNAL        1     27
FT   CHAIN        28    235       EXOTOXIN TYPE C.
FT   CONFLICT     53     53       N -> D (IN REF. 1).
FT   HELIX        33     44
FT   STRAND       49     59
FT   STRAND       63     67
FT   HELIX        69     72
FT   TURN         74     75
FT   STRAND       77     81
FT   HELIX        84     87
FT   TURN         88     89
FT   TURN         92     93
FT   STRAND       95    100
FT   TURN        108    109
FT   STRAND      110    114
FT   STRAND      117    119
FT   STRAND      127    128
FT   STRAND      131    135
FT   TURN        136    137
FT   STRAND      141    142
FT   TURN        144    145
FT   STRAND      148    149
FT   STRAND      153    155
FT   HELIX       156    171
FT   TURN        173    174
FT   TURN        176    177
FT   STRAND      182    189
FT   TURN        190    191
FT   STRAND      194    198
FT   TURN        204    205
FT   HELIX       208    212
FT   HELIX       213    217
FT   STRAND      220    222
FT   HELIX       223    225
FT   STRAND      226    235
SQ   SEQUENCE   235 AA;  27371 MW;  070534ABB952C1E0 CRC64;

  Query Match             17.4%;  Score 205;  DB 1;  Length 235;
  Best Local Similarity   29.5%;  Pred. No. 7.7e-09;
  Matches   67;  Conservative   47;  Mismatches   83;  Indels   30;  Gaps   12;

Qy          7 SQLHRSSLVKNLQNI----YFLYEGDPVTHENVKSVDQLLSHHLIYNVSGPNYD----KL 58
              | : :|   |:: |:     : |   |  ::| : |:   :| |  |:    |      :
Db         22 SPIIKSDSKKDISNVKSDLLYAYTITPYDYKNCR-VNFSTTHTL--NIDTQKYRGKDYYI 78

Qy         59 KTELKNQEMATLFKDKNVDIYGVEYYHLCYLCENAERSACIYGGVT---NHEGNHLEIPK 115
               :|:  :      :| :||::|: :| |     |:     ||||:|   |:: ||
Db         79 SSEMSYEASQKFKRDDHVDVFGL-FYIL-----NSHTGEYIYGGITPAQNNKVNH----- 127

Qy        116 KIVVKVSIDG--IQSLSFDIETNKKMVTAQELDYKVRKYLTDNKQLYTNGPSKYETGYIK 173
              |::  : | |   |:|:  |   | :|| ||:|:|:|||| || ::| :  | | :| |:
Db        128 KLLGNLFISGESQQNLNNKIILEKDIVTFQEIDFKIRKYLMDNYKIY-DATSPYVSGRIE 186

Qy        174 FIPKNKESFWFDFFPEP-EFTQSKYLMIYKDNETLD-SNTSQIEVYL 218
              |: :    | |  | | |:|     ||||  ::  | |  ::||
Db        187 IGTKDGKHEQIDLFDSPNEGTRSDIFAKYKDNRIINMKNFSHFDIYL 233


RESULT 13
SPEG_STRPY
ID   SPEG_STRPY     STANDARD;      PRT;   234 AA.
AC   Q9X5C7;
DT   16-OCT-2001 (Rel. 40, Created)
DT   16-OCT-2001 (Rel. 40, Last sequence update)
DT   16-OCT-2001 (Rel. 40, Last annotation update)
DE   Exotoxin type G precursor (SPE G).
GN   SPEG OR SPY0212.
OS   Streptococcus pyogenes.
OC   Bacteria; Firmicutes; Lactobacillales; Streptococcaceae;
OC   Streptococcus.
OX   NCBI_TaxID=1314;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=M1;
RX   MEDLINE=99093428; PubMed=9874566;
RA   Proft T., Moffatt S.L., Berkahn C.J., Fraser J.D.;
RT   "Identification and characterization of novel superantigens from
RT   Streptococcus pyogenes.";
RL   J. Exp. Med. 189:89-102(1999).
RN   [2]
RP   SEQUENCE FROM N.A.
RC   STRAIN=SF370 / ATCC 700294 / Serotype M1;
RX   MEDLINE=21192684; PubMed=11296296;
RA   Ferretti J.J., McShan W.M., Ajdic D.J., Savic D.J., Savic G., Lyon K.,
RA   Primeaux C., Sezate S., Suvorov A.N., Kenton S., Lai H.S., Lin S.P.,
RA   Qian Y., Jia H.G., Najar F.Z., Ren Q., Zhu H., Song L., White J.,
RA   Yuan X., Clifton S.W., Roe B.A., McLaughlin R.;
RT   "Complete genome sequence of an M1 strain of Streptococcus pyogenes.";
RL   Proc. Natl. Acad. Sci. U.S.A. 98:4658-4663(2001).
CC   -!- FUNCTION: Mitogenic for human peripheral blood lymphocytes.
CC   -!- SUBUNIT: Binds to major histocompatibility complex class II beta
CC       chain.
CC   -!- SIMILARITY: Belongs to the staphylococcal/streptococcal toxin
CC       family.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
```

```
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ------------------------------------------------------------------------
DR   EMBL; AF124499; AAD30988.1; -.
DR   EMBL; AE006489; AAK33303.1; -.
DR   HSSP; P13380; 1AN8.
DR   InterPro; IPR008992; Bact_endotox.
DR   InterPro; IPR006177; Bctrl_tox.
DR   InterPro; IPR006123; Stap/Strep_toxin.
DR   InterPro; IPR006126; Staph/Strept_tox.
DR   InterPro; IPR006173; Staph_tox_OB.
DR   Pfam; PF02876; Stap_Strp_tox_C; 1.
DR   Pfam; PF01123; Stap_Strp_toxin; 1.
DR   PRINTS; PR00279; BACTRLTOXIN.
DR   PROSITE; PS00277; STAPH_STREP_TOXIN_1; FALSE_NEG.
DR   PROSITE; PS00278; STAPH_STREP_TOXIN_2; 1.
KW   Toxin; Signal; Complete proteome.
FT   SIGNAL        1     24       POTENTIAL.
FT   CHAIN        25    234       EXOTOXIN TYPE G.
SQ   SEQUENCE   234 AA;  27262 MW;  49525C49E4BA2052 CRC64;

  Query Match             13.8%;  Score 162;  DB 1;  Length 234;
  Best Local Similarity   24.1%;  Pred. No. 1.5e-05;
  Matches   53;  Conservative   42;  Mismatches   81;  Indels   44;  Gaps   10;

Qy         33 ENVKSVDQLLSHHLIYNVSGPNYDK-----------LKTELKNQEMA----------- 68
              ||:| : : |     ||::  :|:           : |: |  |
Db         26 ENLKDLKRSL--RFAYNITPCDYENVEIAFVTTNSIHINTKQKRSECILYVDSIVSLGIT 83

Qy         69 -TLFKDKNVDIYGVEYYHLCYLCENAERSACIYGGVTNH--EGNHLEIPKKIVVKVSIDG 125
               |   ||::|: |       :|      ||||:  |  :||      : |  :: ||
Db         84 DQFIKGDKVDVFGLPYNFSPPYVDN------IYGGIVKHSNQGNK---SLQFVGILNQDG 134

Qy        126 IQSL--SFDIETNKKMVTAQELDYKVRKYLTDNKQLYTNGPSKYETGYIKFIPKNKESFW 183
              ::   |  :   ||  | || |:|:||:| :   :| :  |:| :| :    |: : :
Db        135 KETYLPSEAVRIKKKQFTLQEFDFKIRKFLMEKYNIY-DSESRYTSGSLFLATKDSKHYE 193

Qy        184 FDFFPEPE--FTQSKYLMIYKDNETLDS-NTSQIEVYLTT 220
              | | : :   ::  :   ||||:  :|   |  ::|| |
Db        194 VDLFNKDDKLLSRDSFFKRYKDNKIFNSEEISHFDIYLKT 233


RESULT 14
TSST_STAAU
ID   TSST_STAAU     STANDARD;      PRT;   234 AA.
AC   P06886;
DT   01-JAN-1988 (Rel. 06, Created)
DT   01-JAN-1988 (Rel. 06, Last sequence update)
DT   15-MAR-2004 (Rel. 43, Last annotation update)
DE   Toxic shock syndrome toxin-1 precursor (TSST-1).
GN   TST.
OS   Staphylococcus aureus.
OC   Bacteria; Firmicutes; Bacillales; Staphylococcus.
OX   NCBI_TaxID=1280;
RN   [1]
RP   SEQUENCE FROM N.A., AND PARTIAL SEQUENCE.
RX   MEDLINE=87057222; PubMed=3782090;
RA   Blomster-Hautamaa D.A., Kreiswirth B.N., Kornblum J.S., Novick R.P.,
RA   Schlievert P.M.;
RT   "The nucleotide and partial amino acid sequence of toxic shock
RT   syndrome toxin-1.";
RL   J. Biol. Chem. 261:15783-15786(1986).
RN   [2]
RP   X-RAY CRYSTALLOGRAPHY (2.5 ANGSTROMS).
RX   MEDLINE=94150598; PubMed=8107781;
RA   Acharya K.R., Passalacqua E.F., Jones E.Y., Harlos K., Stuart D.I.,
RA   Brehm R.D., Tranter H.S.;
RT   "Structural basis of superantigen action inferred from crystal
RT   structure of toxic-shock syndrome toxin-1.";
RL   Nature 367:94-97(1994).
RN   [3]
RP   X-RAY CRYSTALLOGRAPHY (2.5 ANGSTROMS).
RX   MEDLINE=94092653; PubMed=8268150;
RA   Prasad G.S., Earhart C.A., Murray D.L., Novick R.P., Schlivert P.M.,
RA   Ohlendorf D.H.;
RT   "Structure of toxic shock syndrome toxin 1.";
RL   Biochemistry 32:13761-13766(1993).
RN   [4]
RP   X-RAY CRYSTALLOGRAPHY (2.07 ANGSTROMS).
RX   MEDLINE=96319751; PubMed=8759320;
RA   Papgeorgiou A.C., Brehm R.D., Leonidas D.D., Tranter H.S.,
RA   Acharya K.R.;
RT   "The refined crystal structure of toxic shock syndrome toxin-1 at
RT   2.07-A resolution.";
RL   J. Mol. Biol. 260:553-569(1996).
RN   [5]
RP   X-RAY CRYSTALLOGRAPHY (2.05 ANGSTROMS).
RX   MEDLINE=97337442; PubMed=9194182;
RA   Prasad G.S., Radhakrishnan R., Mitchell D.T., Earhart C.A.,
RA   Dinges M.M., Cook W.J., Schlivert P.M., Ohlendorf D.H.;
RT   "Refined structures of three crystal forms of toxic shock syndrome
RT   toxin-1 and of a tetramutant with reduced activity.";
RL   Protein Sci. 6:1220-1227(1997).
RN   [6]
RP   X-RAY CRYSTALLOGRAPHY (1.95 ANGSTROMS) OF MUTANTS.
RX   MEDLINE=98254504; PubMed=9585531;
RA   Earhart C.A., Mitchell D.T., Murray D.L., Pinheiro D.M., Matsumura M.,
RA   Schlievert P.M., Ohlendorf D.H.;
RT   "Structures of five mutants of toxic shock syndrome toxin-1 with
RT   reduced biological activity.";
RL   Biochemistry 37:7194-7202(1998).
CC   -!- FUNCTION: Responsible for the symptoms of toxic shock syndrome.
CC   -!- SUBCELLULAR LOCATION: Secreted.
CC   -!- SIMILARITY: Belongs to the staphylococcal/streptococcal toxin
CC       family.
CC   ------------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ------------------------------------------------------------------------
DR   EMBL; J02615; AAA26682.1; -.
DR   PIR; A24606; XCSAS1.
DR   PDB; 2TSS; 24-DEC-97.
DR   PDB; 3TSS; 24-DEC-97.
DR   PDB; 4TSS; 24-DEC-97.
DR   PDB; 5TSS; 24-DEC-97.
DR   PDB; 1QIL; 12-AUG-97.
DR   PDB; 2QIL; 12-AUG-97.
DR   PDB; 1AW7; 18-NOV-98.
DR   PDB; 1TS2; 16-DEC-98.
DR   PDB; 1TS3; 16-DEC-98.
DR   PDB; 1TS4; 16-DEC-98.
DR   PDB; 1TS5; 16-DEC-98.
DR   InterPro; IPR008992; Bact_endotox.
DR   InterPro; IPR006123; Stap/Strep_toxin.
DR   InterPro; IPR006126; Staph/Strept_tox.
DR   InterPro; IPR008375; Staph_exotoxin.
DR   InterPro; IPR006173; Staph_tox_OB.
DR   InterPro; IPR006125; Staph_toxin.
DR   Pfam; PF02876; Stap_Strp_tox_C; 1.
DR   Pfam; PF01123; Stap_Strp_toxin; 1.
DR   PRINTS; PR01800; STAPHEXOTOXN.
DR   PRINTS; PR01501; TOXICSSTOXIN.
DR   PROSITE; PS00277; STAPH_STREP_TOXIN_1; FALSE_NEG.
DR   PROSITE; PS00278; STAPH_STREP_TOXIN_2; 1.
KW   Toxin; Superantigen; Signal; 3D-structure.
FT   SIGNAL        1     40
FT   CHAIN        41    234       TOXIC SHOCK SYNDROME TOXIN-1.
FT   HELIX        46     54
FT   STRAND       58     69
FT   TURN         70     71
FT   STRAND       72     76
```

```
FT   TURN          78     79
FT   STRAND        82     86
FT   TURN          90     91
FT   TURN          98     99
FT   STRAND       101    108
FT   STRAND       113    114
FT   TURN         116    117
FT   STRAND       120    125
FT   STRAND       128    129
FT   STRAND       133    138
FT   STRAND       142    146
FT   TURN         147    148
FT   STRAND       149    151
FT   STRAND       159    161
FT   STRAND       163    164
FT   HELIX        166    181
FT   TURN         183    185
FT   TURN         187    188
FT   STRAND       192    198
FT   TURN         199    200
FT   STRAND       203    207
FT   TURN         208    209
FT   HELIX        214    216
FT   STRAND       221    222
FT   HELIX        223    225
FT   STRAND       226    234
SQ   SEQUENCE   234 AA;  26306 MW;  E95789FF9A1D7AB4 CRC64;

  Query Match             8.6%;  Score 101;  DB 1;  Length 234;
  Best Local Similarity   22.6%;  Pred. No. 0.65;
  Matches   48;  Conservative   31;  Mismatches   99;  Indels   34;  Gaps    6;

Qy          3 DPDPSQLHRSSLVK--------NLQNIYFLYEGDPVTHENVKSVDQLLSHHLIYNVSG-- 52
              |  |  |  : ::|         |::::   |      |  | : :|  |    | |  |
Db         23 DFTPVPLSSNQIIKTAKASTNDNIKDLLDWYSSGSDTFTNSEVLDNSLGSMRIKNTDGSI 82

Qy         53 -------PNYDKLKTELKNQEMATLFKDKNVDIYGVEYYHLCYLCENAERSACIYGGVTN 105
                     | |    |: :  :: |    |:       | |                ||||
Db         83 SLIIFPSPYYSPAFTKGEKVDLNTKRTKKSQHTSEGTYIHF------------QISGVTN 130

Qy        106 HEGNHLEIPKKIVVKVSIDGIQS-LSFDIETNKKMVTAQELDYKVRKYLTDNKQLYTNGP 164
              |   |  | :: :|| : |  | | :  : :|| :    ||:::|  ||     || :
Db        131 TE--KLPTPIELPLKVKVHGKDSPLKYGPKFDKKQLAISTLDFEIRHQLTQIHGLYRS-- 186

Qy        165 SKYETGYIKFIPKNKESFWFDFFPEPEFTQSK 196
              |    || |    :  ::  |   : |:   |
Db        187 SDKTGGYWKITMNDGSTYQSDLSKKFEYNTEK 218


RESULT 15
DSC1_HUMAN
ID   DSC1_HUMAN      STANDARD;      PRT;   894 AA.
AC   Q08554;
DT   01-NOV-1997 (Rel. 35, Created)
DT   01-NOV-1997 (Rel. 35, Last sequence update)
DT   10-OCT-2003 (Rel. 42, Last annotation update)
DE   Desmocollin 1A/1B precursor (Desmosomal glycoprotein 2/3) (DG2/DG3).
GN   DSC1.
OS   Homo sapiens (Human).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX   NCBI_TaxID=9606;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   TISSUE=Foreskin;
RX   MEDLINE=93283249; PubMed=8507556;
RA   Theis D.G., Koch P.J., Franke W.W.;
RT   "Differential synthesis of type 1 and type 2 desmocollin mRNAs in
RT   human stratified epithelia.";
RL   Int. J. Dev. Biol. 37:101-110(1993).
RN   [2]
RP   SEQUENCE FROM N.A.
RC   TISSUE=Foreskin;
RA   Zimbelmann R.;
RL   Submitted (JUN-1994) to the EMBL/GenBank/DDBJ databases.
RN   [3]
RP   SEQUENCE FROM N.A.
RC   TISSUE=Skin;
RX   MEDLINE=94116981; PubMed=8288219;
RA   King I.A., Arnemann J., Spurr N.K., Buxton R.S.;
RT   "Cloning of the cDNA (DSC1) coding for human type 1 desmocollin and
RT   its assignment to chromosome 18.";
RL   Genomics 18:185-194(1993).
RN   [4]
RP   SEQUENCE OF 135-151 AND 283-292.
RX   MEDLINE=91323543; PubMed=1713860;
RA   King I.A., Magee A.I., Rees D.A., Buxton R.S.;
RT   "Keratinization is associated with the expression of a new protein
RT   related to the desmosomal cadherins DGII/III.";
RL   FEBS Lett. 286:9-12(1991).
CC   -!- FUNCTION: Component of intercellular desmosome junctions. Involved
CC       in the interaction of plaque proteins and intermediate filaments
CC       mediating cell-cell adhesion. May contribute to epidermal cell
CC       positioning (stratification) by mediating differential
CC       adhesiveness between cells that express different isoforms. Linked
CC       to the keratinization of epithelial tissues.
CC   -!- SUBCELLULAR LOCATION: Type I membrane protein.
CC   -!- ALTERNATIVE PRODUCTS:
CC       Event=Alternative splicing; Named isoforms=2;
CC       Name=1A; Synonyms=DG2;
CC         IsoId=Q08554-1; Sequence=Displayed;
CC       Name=1B; Synonyms=DG3;
CC         IsoId=Q08554-2; Sequence=VSP_000651, VSP_000652;
CC   -!- TISSUE SPECIFICITY: Strongly expressed in epidermis, less in lymph
CC       node and tongue.
CC   -!- DOMAIN: Calcium may be bound by the cadherin-like repeats
CC       (Potential).
CC   -!- SIMILARITY: Contains 5 cadherin domains.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; Z34522; CAA84279.1; -.
DR   EMBL; Z34522; CAA84278.1; -.
DR   EMBL; X72925; CAA51428.1; -.
DR   EMBL; X72925; CAA51429.1; -.
DR   PIR; I37281; I37281.
DR   PIR; I37282; I37282.
DR   HSSP; P15116; 1NCJ.
DR   Genew; HGNC:3035; DSC1.
DR   MIM; 125643; -.
DR   GO; GO:0005921; C:gap junction; TAS.
DR   GO; GO:0005624; C:membrane fraction; TAS.
DR   InterPro; IPR002126; Cadherin.
DR   Pfam; PF00028; cadherin; 5.
DR   PRINTS; PR00205; CADHERIN.
DR   SMART; SM00112; CA; 5.
DR   PROSITE; PS00232; CADHERIN_1; 3.
DR   PROSITE; PS50268; CADHERIN_2; 5.
KW   Cell adhesion; Glycoprotein; Transmembrane; Repeat; Signal;
KW   Cytoskeleton; Calcium-binding; Alternative splicing.
FT   SIGNAL        1     29       POTENTIAL.
FT   PROPEP       30    134       POTENTIAL.
FT   CHAIN       135    894       DESMOCOLLIN 1A/1B.
FT   DOMAIN      135    691       EXTRACELLULAR (POTENTIAL).
FT   TRANSMEM    692    714       POTENTIAL.
FT   DOMAIN      715    894       CYTOPLASMIC (POTENTIAL).
FT   DOMAIN      135    242       CADHERIN 1.
FT   DOMAIN      243    354       CADHERIN 2.
FT   DOMAIN      355    471       CADHERIN 3.
FT   DOMAIN      472    575       CADHERIN 4.
```

```
FT   DOMAIN      576    682       CADHERIN 5.
FT   CARBOHYD    165    165       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD    546    546       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   VARSPLIC    830    840       KVYLCGQDEEH -> ESIRGHTLIKN (in isoform
FT                                1B).
FT                                /FTId=VSP_000651.
FT   VARSPLIC    841    894       Missing (in isoform 1B).
FT                                /FTId=VSP_000652.
FT   CONFLICT    132    132       T -> S (IN REF. 3).
SQ   SEQUENCE   894 AA;  100044 MW;  44BA33038699E3E1 CRC64;

  Query Match             8.6%;  Score 101;  DB 1;  Length 894;
  Best Local Similarity   19.8%;  Pred. No. 3.2;
  Matches   51;  Conservative   43;  Mismatches   79;  Indels   84;  Gaps   10;

Qy          4 PDPSQLHRSSLVKNLQNIYFLYEGDPVTHENVKSVDQLLSHHLIYNVSGPNYDKLKTELK 63
              | |: |  :||    |::          : ::| |   :: : |::|||  ||     |
Db        138 PIPASLMENSLGPFPQHV----------QQIQS-DAAQNYTIFYSISGPGVDKEPFNLF 185

Qy         64 NQEMAT--LFKDKNVDIYGVEYYHLCYLCENAERSACIYGGVTNHEGNHLEIPKKIVVKV 121
              |  |  :|  :::|    | :              :||  |  :|   | |  :::|:
Db        186 YIEKDTGDIFCTRSIDREKYEQF------------ALYGYATTADGYAPEYPLPLIIKI 232

Qy        122 SIDGIQSLSFDIE------------TNKKMVTAQELD----------YKVRKYLTDNKQ 158
              |   :  |:                |:  ||| :||          ||: : : |: :
Db        233 EDDNDNAPYFEHRVTIFTVPENCRSGTSVGKVTATDLDEPDTLHTRLKYKILQQIPDHPK 292

Qy        159 LYTNGPSKYETGYI----KFIPKNKESFW------------FDFF-------------- 187
              ::  |   :|| |     |: : |   :              |  |
Db        293 HFSIHP---DTGVITTTTPFLDREKCDTYQLIMEVRDMGGQPFGLFNTGTITISLEDEND 349

Qy        188 PEPEFTQSKYLMIYKDN 204
              | ||:: |:   ::|
Db        350 NPPSFTETSYVTEVEEN 366


Search completed: May  7, 2004, 12:08:08
Job time : 19 secs
```

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: May 7, 2004, 12:04:53 ; Search time 46 Seconds
(without alignments)
1515.859 Million cell updates/sec

Title: US-10-625-221-14
Perfect score: 1178
Sequence: 1 QQDPDPSQLHRSSLVKNLQN.....KDNETLDSNTSQIEVYLTTK 221

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 1017041 seqs, 315518202 residues

Total number of hits satisfying chosen parameters: 1017041

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : SPTREMBL_25:*
1: sp_archea:*
2: sp_bacteria:*
3: sp_fungi:*
4: sp_human:*
5: sp_invertebrate:*
6: sp_mammal:*
7: sp_mhc:*
8: sp_organelle:*
9: sp_phage:*
10: sp_plant:*
11: sp_rodent:*
12: sp_virus:*
13: sp_vertebrate:*
14: sp_unclassified:*
15: sp_rvirus:*
16: sp_bacteriap:*
17: sp_archeap:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	1168	99.2	251	16	Q8K6K5	Q8k6k5 streptococc
2	1134	96.3	236	2	P97163	P97163 streptococc
3	1133	96.2	236	2	Q54779	Q54779 streptococc
4	1128	95.8	236	2	Q57453	Q57453 streptococc
5	1125	95.5	222	2	Q9R931	Q9r931 streptococc
6	1025	87.0	236	2	Q54696	Q54696 streptococc
7	1011	85.8	222	2	Q9S5Z4	Q9s5z4 streptococc
8	1011	85.8	222	2	Q938P4	Q938p4 streptococc
9	549	46.6	239	2	Q53678	Q53678 staphylococ
10	549	46.6	271	2	Q9F0L6	Q9f0l6 staphylococ
11	548	46.5	239	2	Q06532	Q06532 staphylococ
12	547	46.4	239	2	Q06533	Q06533 staphylococ
13	547	46.4	266	16	Q8NXJ6	Q8nxj6 staphylococ
14	546	46.3	239	2	O05157	O05157 staphylococ
15	546	46.3	239	2	Q06531	Q06531 staphylococ
16	542	46.0	239	2	Q06535	Q06535 staphylococ
17	537	45.6	234	2	Q9R5X4	Q9r5x4 staphylococ
18	537	45.6	239	2	Q06534	Q06534 staphylococ
19	517.5	43.9	260	2	Q54971	Q54971 streptococc
20	512.5	43.5	260	2	Q54738	Q54738 streptococc
21	512.5	43.5	260	16	Q54739	Q54739 streptococc
22	509	43.2	259	2	Q936G4	Q936g4 staphylococ
23	468	39.7	233	2	Q8RR77	Q8rr77 staphylococ
24	458	38.9	258	2	Q9EZM3	Q9ezm3 staphylococ
25	455	38.6	258	2	Q9ZNF2	Q9znf2 staphylococ
26	337	28.6	136	16	Q99T49	Q99t49 staphylococ
27	323.5	27.5	260	16	Q99T46	Q99t46 staphylococ
28	323.5	27.5	261	2	Q9EZM8	Q9ezm8 staphylococ
29	315.5	26.8	260	16	Q99SU3	Q99su3 staphylococ
30	297.5	25.3	258	16	Q9EZM4	Q9ezm4 staphylococ
31	293.5	24.9	260	16	Q931M4	Q931m4 staphylococ
32	291.5	24.7	268	2	O85217	O85217 staphylococ
33	285.5	24.2	108	2	Q9EZM5	Q9ezm5 staphylococ
34	284	24.1	240	16	Q9F0L7	Q9f0l7 staphylococ
35	283	24.0	240	16	Q8NXJ5	Q8nxj5 staphylococ
36	281.5	23.9	217	2	Q8RR76	Q8rr76 staphylococ
37	281.5	23.9	241	16	Q53585	Q53585 staphylococ
38	265.5	22.5	239	16	Q99T47	Q99t47 staphylococ
39	264.5	22.5	239	2	Q9EZM7	Q9ezm7 staphylococ
40	246.5	20.9	256	2	Q8VLW7	Q8vlw7 staphylococ
41	246	20.9	229	16	Q8NW97	Q8nw97 staphylococ
42	245.5	20.8	242	2	Q93CC6	Q93cc6 staphylococ
43	243.5	20.7	242	16	Q8NVM3	Q8nvm3 staphylococ
44	237	20.1	242	16	Q8NVM2	Q8nvm2 staphylococ
45	235	19.9	242	2	O54476	O54476 staphylococ

ALIGNMENTS

RESULT 1
Q8K6K5
ID Q8K6K5 PRELIMINARY; PRT; 251 AA.
AC Q8K6K5;
DT 01-OCT-2002 (TrEMBLrel. 22, Created)
DT 01-OCT-2002 (TrEMBLrel. 22, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE Exotoxin type A-phage associated (SpeA precursor).
GN SPEA3 OR SPYM3_1301 OR SPS0560.
OS Streptococcus pyogenes (serotype M3).
OC Bacteria; Firmicutes; Lactobacillales; Streptococcaceae;
OC Streptococcus.
OX NCBI_TaxID=198466;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=MGAS315 / Serotype M3;
RX MEDLINE=22133808; PubMed=12122206;
RA Beres S.B., Sylva G.L., Barbian K.D., Lei B., Hoff J.S.,
RA Mammarella N.D., Liu M.-Y., Smoot J.C., Porcella S.F., Parkins L.D.,
RA Campbell D.S., Smith T.M., McCormick J.K., Leung D.Y.M.,
RA Schlievert P.M., Musser J.M.;
RT "Genome sequence of a serotype M3 strain of group A Streptococcus:
RT phage-encoded toxins, the high-virulence phenotype, and clone
RT emergence.";
RL Proc. Natl. Acad. Sci. U.S.A. 99:10078-10083(2002).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=SSI-1 / Serotype M3;
RA Nakagawa I., Kurokawa K., Nakata M., Tomiyasu Y., Yamashita A.,
RA Yamazaki K., Okahashi N., Kawabata S., Yasunaga T., Hattori M.,
RA Hayashi H., Hamada S.;
RT "The genome of invasive Streptococcus pyogenes; a comparative analysis
RT of S. pyogenes SSI-1, SF370 and MGAS8232.";
RL Submitted (MAY-2002) to the EMBL/GenBank/DDBJ databases.
DR EMBL; AE014161; AAM79908.1; -.
DR EMBL; AP005142; BAC63655.1; -.
DR PIR; A60108; A60108.
DR GO; GO:0005576; C:extracellular; IEA.
DR GO; GO:0015070; F:toxin activity; IEA.

```
DR   GO; GO:0009405; P:pathogenesis; IEA.
DR   InterPro; IPR008992; Bact_endotox.
DR   InterPro; IPR006177; Bctrl_tox.
DR   InterPro; IPR006123; Stap/Strep_toxin.
DR   InterPro; IPR006126; Staph/Strept_tox.
DR   InterPro; IPR006173; Staph_tox_OB.
DR   Pfam; PF01123; Stap_Strp_toxin; 1.
DR   Pfam; PF02876; Stap_Strp_tox_C; 1.
DR   PRINTS; PR00279; BACTRLTOXIN.
DR   PROSITE; PS00277; STAPH_STREP_TOXIN_1; 1.
DR   PROSITE; PS00278; STAPH_STREP_TOXIN_2; 1.
KW   Complete proteome.
SQ   SEQUENCE   251 AA;  29260 MW;  05E782CDA01BFCD5 CRC64;

  Query Match             99.2%;  Score 1168;  DB 16;  Length 251;
  Best Local Similarity   99.1%;  Pred. No. 5e-78;
  Matches  219;  Conservative    1;  Mismatches    1;  Indels    0;  Gaps    0;

Qy          1 QQDPDPSQLHRSSLVKNLQNIYFLYEGDPVTHENVKSVDQLLSHHLIYNVSGPNYDKLKT 60
              ||||||||||||||||||||||||||||||||||||||||||| ||||||||||||||||
Db         31 QQDPDPSQLHRSSLVKNLQNIYFLYEGDPVTHENVKSVDQLLSHDLIYNVSGPNYDKLKT 90

Qy         61 ELKNQEMATLFKDKNVDIYGVEYYHLCYLCENAERSACIYGGVTNHEGNHLEIPKKIVVK 120
              ||||||||||||||:|||||||||||||||||||||||||||||||||||||||||||||
Db         91 ELKNQEMATLFKDKNIDIYGVEYYHLCYLCENAERSACIYGGVTNHEGNHLEIPKKIVVK 150

Qy        121 VSIDGIQSLSFDIETNKKMVTAQELDYKVRKYLTDNKQLYTNGPSKYETGYIKFIPKNKE 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        151 VSIDGIQSLSFDIETNKKMVTAQELDYKVRKYLTDNKQLYTNGPSKYETGYIKFIPKNKE 210

Qy        181 SFWFDFFPEPEFTQSKYLMIYKDNETLDSNTSQIEVYLTTK 221
              |||||||||||||||||||||||||||||||||||||||||
Db        211 SFWFDFFPEPEFTQSKYLMIYKDNETLDSNTSQIEVYLTTK 251


RESULT 2
P97163
ID   P97163      PRELIMINARY;      PRT;   236 AA.
AC   P97163; P97164;
DT   01-MAY-1997 (TrEMBLrel. 03, Created)
DT   01-MAY-1997 (TrEMBLrel. 03, Last sequence update)
DT   01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE   Type A exotoxin precursor (Fragment).
GN   SPEA.
OS   Streptococcus pyogenes.
OC   Bacteria; Firmicutes; Lactobacillales; Streptococcaceae;
OC   Streptococcus.
OX   NCBI_TaxID=1314;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=MGAS156, and MGAS500;
RX   MEDLINE=92044323; PubMed=1940804;
RA   Nelson K., Schlievert P.M., Selander R.K., Musser J.M.;
RT   "Characterization and clonal distribution of four alleles of the speA
RT   gene encoding pyrogenic exotoxin A (scarlet fever toxin) in
RT   Streptococcus pyogenes.";
RL   J. Exp. Med. 174:1271-1274(1991).
DR   EMBL; X61556; CAA43754.1; -.
DR   EMBL; X61557; CAA43755.1; -.
DR   EMBL; X61560; CAA43758.1; -.
DR   EMBL; X61555; CAA43753.1; -.
DR   EMBL; X61558; CAA43756.1; -.
DR   EMBL; X61559; CAA43757.1; -.
DR   EMBL; X61554; CAA43752.1; -.
DR   PIR; A60108; A60108.
DR   HSSP; P08095; 1B1Z.
DR   GO; GO:0005576; C:extracellular; IEA.
DR   GO; GO:0015070; F:toxin activity; IEA.
DR   GO; GO:0009405; P:pathogenesis; IEA.
DR   InterPro; IPR008992; Bact_endotox.
DR   InterPro; IPR006177; Bctrl_tox.
DR   InterPro; IPR006123; Stap/Strep_toxin.
DR   InterPro; IPR006126; Staph/Strept_tox.
DR   InterPro; IPR006173; Staph_tox_OB.
DR   Pfam; PF01123; Stap_Strp_toxin; 1.
DR   Pfam; PF02876; Stap_Strp_tox_C; 1.
DR   PRINTS; PR00279; BACTRLTOXIN.
DR   PROSITE; PS00277; STAPH_STREP_TOXIN_1; 1.
DR   PROSITE; PS00278; STAPH_STREP_TOXIN_2; 1.
KW   Signal.
FT   NON_TER       1      1
FT   SIGNAL       <1     22       POTENTIAL.
FT   CHAIN         23   >236       TYPE A EXOTOXIN.
FT   NON_TER      236    236
SQ   SEQUENCE   236 AA;  27454 MW;  3FB3F41ABDC13A84 CRC64;

  Query Match             96.3%;  Score 1134;  DB 2;  Length 236;
  Best Local Similarity   99.5%;  Pred. No. 1.5e-75;
  Matches  213;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy          1 QQDPDPSQLHRSSLVKNLQNIYFLYEGDPVTHENVKSVDQLLSHHLIYNVSGPNYDKLKT 60
              |||||||||||||||||||||||||||||||||||||||||| ||||||||||||||||
Db         23 QQDPDPSQLHRSSLVKNLQNIYFLYEGDPVTHENVKSVDQLLSHDLIYNVSGPNYDKLKT 82

Qy         61 ELKNQEMATLFKDKNVDIYGVEYYHLCYLCENAERSACIYGGVTNHEGNHLEIPKKIVVK 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         83 ELKNQEMATLFKDKNVDIYGVEYYHLCYLCENAERSACIYGGVTNHEGNHLEIPKKIVVK 142

Qy        121 VSIDGIQSLSFDIETNKKMVTAQELDYKVRKYLTDNKQLYTNGPSKYETGYIKFIPKNKE 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        143 VSIDGIQSLSFDIETNKKMVTAQELDYKVRKYLTDNKQLYTNGPSKYETGYIKFIPKNKE 202

Qy        181 SFWFDFFPEPEFTQSKYLMIYKDNETLDSNTSQI 214
              |||||||||||||||||||||||||||||||||
Db        203 SFWFDFFPEPEFTQSKYLMIYKDNETLDSNTSQI 236


RESULT 3
Q54779
ID   Q54779      PRELIMINARY;      PRT;   236 AA.
AC   Q54779; Q54613; Q54736; Q54740; Q54741;
DT   01-NOV-1996 (TrEMBLrel. 01, Created)
DT   01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
DT   01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE   Type A exotoxin precursor (Fragment).
GN   SPEA.
OS   Streptococcus pyogenes.
OC   Bacteria; Firmicutes; Lactobacillales; Streptococcaceae;
OC   Streptococcus.
OX   NCBI_TaxID=1314;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=MGAS624 AND MGAS158 AND MGAS485 AND MGAS491, and MGAS495;
RX   MEDLINE=92044323; PubMed=1940804;
RA   Nelson K., Schlievert P.M., Selander R.K., Musser J.M.;
RT   "Characterization and clonal distribution of four alleles of the speA
RT   gene encoding pyrogenic exotoxin A (scarlet fever toxin) in
RT   Streptococcus pyogenes.";
RL   J. Exp. Med. 174:1271-1274(1991).
DR   EMBL; X61569; CAA43767.1; -.
DR   EMBL; X61572; CAA43770.1; -.
DR   EMBL; X61568; CAA43766.1; -.
DR   EMBL; X61570; CAA43768.1; -.
DR   EMBL; X61571; CAA43769.1; -.
DR   PIR; A60108; A60108.
DR   HSSP; P08095; 1B1Z.
DR   GO; GO:0005576; C:extracellular; IEA.
DR   GO; GO:0015070; F:toxin activity; IEA.
DR   GO; GO:0009405; P:pathogenesis; IEA.
DR   InterPro; IPR008992; Bact_endotox.
DR   InterPro; IPR006177; Bctrl_tox.
DR   InterPro; IPR006123; Stap/Strep_toxin.
DR   InterPro; IPR006126; Staph/Strept_tox.
DR   InterPro; IPR006173; Staph_tox_OB.
DR   Pfam; PF01123; Stap_Strp_toxin; 1.
DR   Pfam; PF02876; Stap_Strp_tox_C; 1.
```

```
DR   PRINTS; PR00279; BACTRLTOXIN.
DR   PROSITE; PS00277; STAPH_STREP_TOXIN_1; 1.
DR   PROSITE; PS00278; STAPH_STREP_TOXIN_2; 1.
KW   Signal.
FT   NON_TER       1      1
FT   SIGNAL       <1     22       POTENTIAL.
FT   CHAIN        23   >236       TYPE A EXOTOXIN.
FT   NON_TER     236    236
SQ   SEQUENCE   236 AA;  27468 MW;  29DF2AD575623A84 CRC64;

  Query Match             96.2%;  Score 1133;  DB 2;  Length 236;
  Best Local Similarity   99.1%;  Pred. No. 1.7e-75;
  Matches  212;  Conservative    1;  Mismatches    1;  Indels    0;  Gaps    0;

Qy          1 QQDPDPSQLHRSSLVKNLQNIYFLYEGDPVTHENVKSVDQLLSHHLIYNVSGPNYDKLKT 60
              ||||||||||||||||||||||||||||||||||||||||||| ||||||||||||||||
Db         23 QQDPDPSQLHRSSLVKNLQNIYFLYEGDPVTHENVKSVDQLLSHDLIYNVSGPNYDKLKT 82

Qy         61 ELKNQEMATLFKDKNVDIYGVEYYHLCYLCENAERSACIYGGVTNHEGNHLEIPKKIVVK 120
              |||||||||||||||:||||||||||||||||||||||||||||||||||||||||||||
Db         83 ELKNQEMATLFKDKNIDIYGVEYYHLCYLCENAERSACIYGGVTNHEGNHLEIPKKIVVK 142

Qy        121 VSIDGIQSLSFDIETNKKMVTAQELDYKVRKYLTDNKQLYTNGPSKYETGYIKFIPKNKE 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        143 VSIDGIQSLSFDIETNKKMVTAQELDYKVRKYLTDNKQLYTNGPSKYETGYIKFIPKNKE 202

Qy        181 SFWFDFFPEPEFTQSKYLMIYKDNETLDSNTSQI 214
              |||||||||||||||||||||||||||||||||
Db        203 SFWFDFFPEPEFTQSKYLMIYKDNETLDSNTSQI 236


RESULT 4
Q57453
ID   Q57453      PRELIMINARY;      PRT;   236 AA.
AC   Q57453;
DT   01-NOV-1996 (TrEMBLrel. 01, Created)
DT   01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
DT   01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE   Type A exotoxin precursor (Fragment).
GN   SPEA.
OS   Streptococcus pyogenes.
OC   Bacteria; Firmicutes; Lactobacillales; Streptococcaceae;
OC   Streptococcus.
OX   NCBI_TaxID=1314;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=MGAS496;
RX   MEDLINE=92044323; PubMed=1940804;
RA   Nelson K., Schlievert P.M., Selander R.K., Musser J.M.;
RT   "Characterization and clonal distribution of four alleles of the speA
RT   gene encoding pyrogenic exotoxin A (scarlet fever toxin) in
RT   Streptococcus pyogenes.";
RL   J. Exp. Med. 174:1271-1274(1991).
DR   EMBL; X61562; CAA43760.1; -.
DR   EMBL; X61563; CAA43761.1; -.
DR   EMBL; X61567; CAA43765.1; -.
DR   EMBL; X61561; CAA43759.1; -.
DR   EMBL; X61564; CAA43762.1; -.
DR   EMBL; X61565; CAA43763.1; -.
DR   EMBL; X61566; CAA43764.1; -.
DR   PIR; A60108; A60108.
DR   HSSP; P08095; 1B1Z.
DR   GO; GO:0005576; C:extracellular; IEA.
DR   GO; GO:0015070; F:toxin activity; IEA.
DR   GO; GO:0009405; P:pathogenesis; IEA.
DR   InterPro; IPR008992; Bact_endotox.
DR   InterPro; IPR006177; Bctrl_tox.
DR   InterPro; IPR006123; Stap/Strep_toxin.
DR   InterPro; IPR006126; Staph/Strept_tox.
DR   InterPro; IPR006173; Staph_tox_OB.
DR   Pfam; PF01123; Stap_Strp_toxin; 1.
DR   Pfam; PF02876; Stap_Strp_tox_C; 1.
DR   PRINTS; PR00279; BACTRLTOXIN.
DR   PROSITE; PS00277; STAPH_STREP_TOXIN_1; 1.
DR   PROSITE; PS00278; STAPH_STREP_TOXIN_2; 1.
KW   Signal.
FT   NON_TER       1      1
FT   SIGNAL       <1     22       POTENTIAL.
FT   CHAIN        23   >236       TYPE A EXOTOXIN.
FT   NON_TER     236    236
SQ   SEQUENCE   236 AA;  27484 MW;  2EF7F41AAC853600 CRC64;

  Query Match             95.8%;  Score 1128;  DB 2;  Length 236;
  Best Local Similarity   99.1%;  Pred. No. 4e-75;
  Matches  212;  Conservative    0;  Mismatches    2;  Indels    0;  Gaps    0;

Qy          1 QQDPDPSQLHRSSLVKNLQNIYFLYEGDPVTHENVKSVDQLLSHHLIYNVSGPNYDKLKT 60
              ||||||||||||||||||||||||||||||||||||||||||| ||||||||||||||||
Db         23 QQDPDPSQLHRSSLVKNLQNIYFLYEGDPVTHENVKSVDQLLSHDLIYNVSGPNYDKLKT 82

Qy         61 ELKNQEMATLFKDKNVDIYGVEYYHLCYLCENAERSACIYGGVTNHEGNHLEIPKKIVVK 120
              |||||||||||||||||| ||||||||||||||||||||||||||||||||||||||||
Db         83 ELKNQEMATLFKDKNVDIYSVEYYHLCYLCENAERSACIYGGVTNHEGNHLEIPKKIVVK 142

Qy        121 VSIDGIQSLSFDIETNKKMVTAQELDYKVRKYLTDNKQLYTNGPSKYETGYIKFIPKNKE 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        143 VSIDGIQSLSFDIETNKKMVTAQELDYKVRKYLTDNKQLYTNGPSKYETGYIKFIPKNKE 202

Qy        181 SFWFDFFPEPEFTQSKYLMIYKDNETLDSNTSQI 214
              |||||||||||||||||||||||||||||||||
Db        203 SFWFDFFPEPEFTQSKYLMIYKDNETLDSNTSQI 236


RESULT 5
Q9R931
ID   Q9R931      PRELIMINARY;      PRT;   222 AA.
AC   Q9R931;
DT   01-MAY-2000 (TrEMBLrel. 13, Created)
DT   01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT   01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE   Exotoxin A (Fragment).
GN   SPEA.
OS   Streptococcus pyogenes.
OC   Bacteria; Firmicutes; Lactobacillales; Streptococcaceae;
OC   Streptococcus.
OX   NCBI_TaxID=1314;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=D709;
RX   MEDLINE=99137798; PubMed=9952369;
RA   Bessen D.E., Izzo M.W., Fiorentino T.R., Caringal R.M.,
RA   Hollingshead S.K., Beall B.;
RT   "Genetic linkage of exotoxin alleles and emm gene markers for tissue
RT   tropism in group A streptococci.";
RL   J. Infect. Dis. 179:627-636(1999).
DR   EMBL; AF055698; AAD11624.1; -.
DR   PIR; A60108; A60108.
DR   HSSP; P08095; 1B1Z.
DR   GO; GO:0005576; C:extracellular; IEA.
DR   GO; GO:0015070; F:toxin activity; IEA.
DR   GO; GO:0009405; P:pathogenesis; IEA.
DR   InterPro; IPR008992; Bact_endotox.
DR   InterPro; IPR006177; Bctrl_tox.
DR   InterPro; IPR006123; Stap/Strep_toxin.
DR   InterPro; IPR006126; Staph/Strept_tox.
DR   InterPro; IPR006173; Staph_tox_OB.
DR   Pfam; PF01123; Stap_Strp_toxin; 1.
DR   Pfam; PF02876; Stap_Strp_tox_C; 1.
DR   PRINTS; PR00279; BACTRLTOXIN.
DR   PROSITE; PS00277; STAPH_STREP_TOXIN_1; 1.
DR   PROSITE; PS00278; STAPH_STREP_TOXIN_2; 1.
FT   NON_TER       1      1
FT   NON_TER     222    222
SQ   SEQUENCE   222 AA;  25759 MW;  48BB7ADDCD91FBA3 CRC64;

  Query Match             95.5%;  Score 1125;  DB 2;  Length 222;
```

```
  Best Local Similarity   99.5%;  Pred. No. 6.2e-75;
  Matches  211;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy          1 QQDPDPSQLHRSSLVKNLQNIYFLYEGDPVTHENVKSVDQLLSHHLIYNVSGPNYDKLKT 60
              ||||||||||||||||||||||||||||||||||||||||||| ||||||||||||||
Db         11 QQDPDPSQLHRSSLVKNLQNIYFLYEGDPVTHENVKSVDQLLSHDLIYNVSGPNYDKLKT 70

Qy         61 ELKNQEMATLFKDKNVDIYGVEYYHLCYLCENAERSACIYGGVTNHEGNHLEIPKKIVVK 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         71 ELKNQEMATLFKDKNVDIYGVEYYHLCYLCENAERSACIYGGVTNHEGNHLEIPKKIVVK 130

Qy        121 VSIDGIQSLSFDIETNKKMVTAQELDYKVRKYLTDNKQLYTNGPSKYETGYIKFIPKNKE 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        131 VSIDGIQSLSFDIETNKKMVTAQELDYKVRKYLTDNKQLYTNGPSKYETGYIKFIPKNKE 190

Qy        181 SFWFDFFPEPEFTQSKYLMIYKDNETLDSNTS 212
              ||||||||||||||||||||||||||||||||
Db        191 SFWFDFFPEPEFTQSKYLMIYKDNETLDSNTS 222


RESULT 6
Q54696
ID   Q54696      PRELIMINARY;      PRT;   236 AA.
AC   Q54696;
DT   01-NOV-1996 (TrEMBLrel. 01, Created)
DT   01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
DT   01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE   Type A exotoxin precursor (Fragment).
GN   SPEA.
OS   Streptococcus pyogenes.
OC   Bacteria; Firmicutes; Lactobacillales; Streptococcaceae;
OC   Streptococcus.
OX   NCBI_TaxID=1314;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=MGAS156;
RX   MEDLINE=92044323; PubMed=1940804;
RA   Nelson K., Schlievert P.M., Selander R.K., Musser J.M.;
RT   "Characterization and clonal distribution of four alleles of the speA
RT   gene encoding pyrogenic exotoxin A (scarlet fever toxin) in
RT   Streptococcus pyogenes.";
RL   J. Exp. Med. 174:1271-1274(1991).
DR   EMBL; X61573; CAA43771.1; -.
DR   PIR; S18789; S18789.
DR   HSSP; P08095; 1B1Z.
DR   GO; GO:0005576; C:extracellular; IEA.
DR   GO; GO:0015070; F:toxin activity; IEA.
DR   GO; GO:0009405; P:pathogenesis; IEA.
DR   InterPro; IPR008992; Bact_endotox.
DR   InterPro; IPR006177; Bctrl_tox.
DR   InterPro; IPR006123; Stap/Strep_toxin.
DR   InterPro; IPR006126; Staph/Strept_tox.
DR   InterPro; IPR006173; Staph_tox_OB.
DR   Pfam; PF01123; Stap_Strp_toxin; 1.
DR   Pfam; PF02876; Stap_Strp_tox_C; 1.
DR   PRINTS; PR00279; BACTRLTOXIN.
DR   PROSITE; PS00277; STAPH_STREP_TOXIN_1; 1.
DR   PROSITE; PS00278; STAPH_STREP_TOXIN_2; 1.
KW   Signal.
FT   NON_TER       1      1
FT   SIGNAL       <1     22       POTENTIAL.
FT   CHAIN        23    >236      TYPE A EXOTOXIN.
FT   NON_TER     236    236
SQ   SEQUENCE   236 AA;  27575 MW;  70F54120E79127DF CRC64;

  Query Match             87.0%;  Score 1025;  DB 2;  Length 236;
  Best Local Similarity   89.7%;  Pred. No. 1.4e-67;
  Matches  192;  Conservative   12;  Mismatches   10;  Indels    0;  Gaps    0;

Qy          1 QQDPDPSQLHRSSLVKNLQNIYFLYEGDPVTHENVKSVDQLLSHHLIYNVSGPNYDKLKT 60
              ||||:|||||||||||||||||||||||| |||||||||||| ||||||| |||||||
Db         23 QQDPNPSQLHRSSLVKNLQNIYFLYEGDPVVHENVKSVDQLLSHDLIYNVSGLNYDKLKT 82

Qy         61 ELKNQEMATLFKDKNVDIYGVEYYHLCYLCENAERSACIYGGVTNHEGNHLEIPKKIVVK 120
              ||||:||:||||:|||||||||||: |||| ||:| |||||||||||||||||| |:||
Db         83 ELKNREMSTLFKNKNVDIYGVEYYYHCYLCRNAKRRACIYGGVTNHEGNHLEIPKNILVK 142

Qy        121 VSIDGIQSLSFDIETNKKMVTAQELDYKVRKYLTDNKQLYTNGPSKYETGYIKFIPKNKE 180
              ||||||||||||||:|||||||||||||:|||||||||||||||||||||||| |:||
Db        143 VSIDGIQSLSFDIETSKKMVTAQELDYKVRKHLTDNKQLYTNGPSKYETGYIKFISKDKE 202

Qy        181 SFWFDFFPEPEFTQSKYLMIYKDNETLDSNTSQI 214
              :|||||||||| | ||||||||||||||:||||
Db        203 TFWFDFFPEPEFNQVKYLMIYKDNETLDSSTSQI 236


RESULT 7
Q9S5Z4
ID   Q9S5Z4      PRELIMINARY;      PRT;   222 AA.
AC   Q9S5Z4;
DT   01-MAY-2000 (TrEMBLrel. 13, Created)
DT   01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT   01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE   Exotoxin type A (Fragment).
GN   SPEA.
OS   Streptococcus pyogenes.
OC   Bacteria; Firmicutes; Lactobacillales; Streptococcaceae;
OC   Streptococcus.
OX   NCBI_TaxID=1314;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=D633;
RX   MEDLINE=99137798; PubMed=9952369;
RA   Bessen D.E., Izzo M.W., Fiorentino T.R., Caringal R.M.,
RA   Hollingshead S.K., Beall B.;
RT   "Genetic linkage of exotoxin alleles and emm gene markers for tissue
RT   tropism in group A streptococci.";
RL   J. Infect. Dis. 179:627-636(1999).
DR   EMBL; AF029051; AAD21315.1; -.
DR   HSSP; P08095; 1B1Z.
DR   GO; GO:0005576; C:extracellular; IEA.
DR   GO; GO:0015070; F:toxin activity; IEA.
DR   GO; GO:0009405; P:pathogenesis; IEA.
DR   InterPro; IPR008992; Bact_endotox.
DR   InterPro; IPR006177; Bctrl_tox.
DR   InterPro; IPR006123; Stap/Strep_toxin.
DR   InterPro; IPR006126; Staph/Strept_tox.
DR   InterPro; IPR006173; Staph_tox_OB.
DR   Pfam; PF01123; Stap_Strp_toxin; 1.
DR   Pfam; PF02876; Stap_Strp_tox_C; 1.
DR   PRINTS; PR00279; BACTRLTOXIN.
DR   PROSITE; PS00277; STAPH_STREP_TOXIN_1; 1.
DR   PROSITE; PS00278; STAPH_STREP_TOXIN_2; 1.
FT   NON_TER       1      1
FT   NON_TER     222    222
SQ   SEQUENCE   222 AA;  25884 MW;  121F8460992818F8 CRC64;

  Query Match             85.8%;  Score 1011;  DB 2;  Length 222;
  Best Local Similarity   89.2%;  Pred. No. 1.4e-66;
  Matches  189;  Conservative   12;  Mismatches   11;  Indels    0;  Gaps    0;

Qy          1 QQDPDPSQLHRSSLVKNLQNIYFLYEGDPVTHENVKSVDQLLSHHLIYNVSGPNYDKLKT 60
              ||||:|||||||||||||||||||||||| |||||||||||| ||||||| |||||||
Db         11 QQDPNPSQLHRSSLVKNLQNIYFLYEGDPVVHENVKSVDQLLSHDLIYNVSGLNYDKLKT 70

Qy         61 ELKNQEMATLFKDKNVDIYGVEYYHLCYLCENAERSACIYGGVTNHEGNHLEIPKKIVVK 120
              ||||:||:||||:|||||||||||: |||| ||:| |||||||||||||||||| |:||
Db         71 ELKNREMSTLFKNKNVDIYGVEYYYHCYLCRNAKRRACIYGGVTNHEGNHLEIPKNILVK 130

Qy        121 VSIDGIQSLSFDIETNKKMVTAQELDYKVRKYLTDNKQLYTNGPSKYETGYIKFIPKNKE 180
              ||||||||||||||:|||||||||||||:|||| |||||||||||||||| |:||
Db        131 VSIDGIQSLSFDIETSKKMVTAQELDYKVRKHLTDNNQLYTNGPSKYETGYIKFISKDKE 190

Qy        181 SFWFDFFPEPEFTQSKYLMIYKDNETLDSNTS 212
              :|||||||||| | ||||||||||||||:||
Db        191 TFWFDFFPEPEFNQVKYLMIYKDNETLDSSTS 222
```

```
RESULT 8
Q938P4
ID   Q938P4      PRELIMINARY;      PRT;   222 AA.
AC   Q938P4;
DT   01-DEC-2001 (TrEMBLrel. 19, Created)
DT   01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
DT   01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE   Pyrogenic exotoxin A (Fragment).
GN   SPEA.
OS   Streptococcus equisimilis.
OC   Bacteria; Firmicutes; Lactobacillales; Streptococcaceae;
OC   Streptococcus.
OX   NCBI_TaxID=119602;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=4951;
RA   Kalia A., Bessen D.E.;
RT   "Presence of streptococcal pyrogenic exotoxin A and C genes in human
RT   isolates of group G Streptococci.";
RL   Submitted (AUG-2001) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; AY049745; AAL06068.1; -.
DR   GO; GO:0005576; C:extracellular; IEA.
DR   GO; GO:0015070; F:toxin activity; IEA.
DR   GO; GO:0009405; P:pathogenesis; IEA.
DR   InterPro; IPR008992; Bact_endotox.
DR   InterPro; IPR006177; Bctrl_tox.
DR   InterPro; IPR006123; Stap/Strep_toxin.
DR   InterPro; IPR006126; Staph/Strept_tox.
DR   InterPro; IPR006173; Staph_tox_OB.
DR   Pfam; PF01123; Stap_Strp_toxin; 1.
DR   Pfam; PF02876; Stap_Strp_tox_C; 1.
DR   PRINTS; PR00279; BACTRLTOXIN.
DR   PROSITE; PS00277; STAPH_STREP_TOXIN_1; 1.
DR   PROSITE; PS00278; STAPH_STREP_TOXIN_2; 1.
FT   NON_TER       1      1
FT   NON_TER     222    222
SQ   SEQUENCE   222 AA;  25884 MW;  121F8460992818F8 CRC64;

  Query Match             85.8%;  Score 1011;  DB 2;  Length 222;
  Best Local Similarity   89.2%;  Pred. No. 1.4e-66;
  Matches  189;  Conservative   12;  Mismatches   11;  Indels    0;  Gaps    0;

Qy          1 QQDPDPSQLHRSSLVKNLQNIYFLYEGDPVTHENVKSVDQLLSHHLIYNVSGPNYDKLKT 60
              ||||:|||||||||||||||||||||||||| |||||||||||| ||||||| |||||||
Db         11 QQDPNPSQLHRSSLVKNLQNIYFLYEGDPVVHENVKSVDQLLSHDLIYNVSGLNYDKLKT 70

Qy         61 ELKNQEMATLFKDKNVDIYGVEYYHLCYLCENAERSACIYGGVTNHEGNHLEIPKKIVVK 120
              ||||:||:||||:||||||||||: |||| ||:| |||||||||||||||||| |:||
Db         71 ELKNREMSTLFKNKNVDIYGVEYYYHCYLCRNAKRRACIYGGVTNHEGNHLEIPKNILVK 130

Qy        121 VSIDGIQSLSFDIETNKKMVTAQELDYKVRKYLTDNKQLYTNGPSKYETGYIKFIPKNKE 180
              ||||||||||||||:|||||||||||||||:|||| |||||||||||||||||| |:||
Db        131 VSIDGIQSLSFDIETSKKMVTAQELDYKVRKHLTDNNQLYTNGPSKYETGYIKFISKDKE 190

Qy        181 SFWFDFFPEPEFTQSKYLMIYKDNETLDSNTS 212
              :|||||||||| | ||||||||||||||:||
Db        191 TFWFDFFPEPEFNQVKYLMIYKDNETLDSSTS 222


RESULT 9
Q53678
ID   Q53678      PRELIMINARY;      PRT;   239 AA.
AC   Q53678;
DT   01-NOV-1996 (TrEMBLrel. 01, Created)
DT   01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
DT   01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE   Enterotoxin (Fragment).
OS   Staphylococcus aureus.
OC   Bacteria; Firmicutes; Bacillales; Staphylococcus.
OX   NCBI_TaxID=1280;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=94011313; PubMed=8406814;
RA   Marr J.C., Lyon J.D., Roberson J.R., Lupher M., Bohach G.A.;
RT   "Characterization of novel type C staphylococcal enterotoxins:
RT   biological and evolutionary implications.";
RL   Infect. Immun. 61:4254-4262(1993).
DR   EMBL; L13376; AAA26620.1; -.
DR   HSSP; P34071; 1SE2.
DR   GO; GO:0005576; C:extracellular; IEA.
DR   GO; GO:0015070; F:toxin activity; IEA.
DR   GO; GO:0009405; P:pathogenesis; IEA.
DR   InterPro; IPR008992; Bact_endotox.
DR   InterPro; IPR006177; Bctrl_tox.
DR   InterPro; IPR006123; Stap/Strep_toxin.
DR   InterPro; IPR006126; Staph/Strept_tox.
DR   InterPro; IPR006173; Staph_tox_OB.
DR   Pfam; PF01123; Stap_Strp_toxin; 1.
DR   Pfam; PF02876; Stap_Strp_tox_C; 1.
DR   PRINTS; PR00279; BACTRLTOXIN.
DR   PROSITE; PS00277; STAPH_STREP_TOXIN_1; 1.
DR   PROSITE; PS00278; STAPH_STREP_TOXIN_2; 1.
FT   NON_TER       1      1
SQ   SEQUENCE   239 AA;  27618 MW;  A13E7EB25C6989C2 CRC64;

  Query Match             46.6%;  Score 549;  DB 2;  Length 239;
  Best Local Similarity   47.2%;  Pred. No. 1.1e-32;
  Matches  111;  Conservative   44;  Mismatches   66;  Indels   14;  Gaps    6;

Qy          1 QQDPDPSQLHRSSLVKNL-QNIYFLYEGDPVTHENVKSVDQLLSHHLIYNVSG---PNYD 56
              | || | :||::|    | :|:  ||:   |:  |||||: |:| ||||:|     |||
Db          3 QPDPTPDELHKASKFTGLMENMKVLYDDRYVSATKVKSVDKFLAHDLIYNISDKKLKNYD 62

Qy         57 KLKTELKNQEMATLFKDKNVDIYGVEYYHLCYLC--ENAER----SACIYGGVTNHEGNH 110
              |:|||| |:::|  :||: ||:||  ||  ||    :|  :      |:|||:| |||||
Db         63 KVKTELLNEDLAKKYKDEVVDVYGSNYYVNCYFSSKDNVGKVTGGKTCMYGGITKHEGNH 122

Qy        111 LEIPK--KIVVKVSIDGIQSLSFDIETNKKMVTAQELDYKVRKYLTDNKQLYTNGPSKYE 168
              :  |   ::::|  :   ::||:::|:|| ||||||| | | :| : | ||    | ||
Db        123 FDNGKLQNVLIRVYENKRNTISFEVQTDKKSVTAQELDIKARNFLINKKNLYEFNSSPYE 182

Qy        169 TGYIKFIPKNKESFWFDFFPEP--EFTQSKYLMIYKDNETLDSNTSQIEVYLTTK 221
              |||||||  | :||:|  | |  :| ||||||:| ||:|:|| : :|||:||||
Db        183 TGYIKFIENNGNTFWYDMMPAPGDKFDQSKYLMMYNDNKTVDSKSVKIEVHLTTK 237


RESULT 10
Q9F0L6
ID   Q9F0L6      PRELIMINARY;      PRT;   271 AA.
AC   Q9F0L6;
DT   01-MAR-2001 (TrEMBLrel. 16, Created)
DT   01-MAR-2001 (TrEMBLrel. 16, Last sequence update)
DT   01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE   Staphylococcal enterotoxin C-bovine.
GN   SEC-BOV.
OS   Staphylococcus aureus.
OC   Bacteria; Firmicutes; Bacillales; Staphylococcus.
OX   NCBI_TaxID=1280;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=20566668; PubMed=11114901;
RA   Fitzgerald J.R., Monday S.R., Foster T.J., Bohach G.A., Hartigan P.J.,
RA   Meaney W.J., Smyth C.J.;
RT   "Characterization of a putative pathogenicity island from bovine
RT   Staphylococcus aureus encoding multiple superantigens.";
RL   J. Bacteriol. 183:63-70(2001).
DR   EMBL; AF217235; AAG29599.1; -.
DR   HSSP; P34071; 1SE2.
DR   GO; GO:0005576; C:extracellular; IEA.
DR   GO; GO:0015070; F:toxin activity; IEA.
DR   GO; GO:0009405; P:pathogenesis; IEA.
DR   InterPro; IPR008992; Bact_endotox.
DR   InterPro; IPR006177; Bctrl_tox.
DR   InterPro; IPR006123; Stap/Strep_toxin.
```

```
DR   InterPro; IPR006126; Staph/Strept_tox.
DR   InterPro; IPR006173; Staph_tox_OB.
DR   Pfam; PF01123; Stap_Strp_toxin; 1.
DR   Pfam; PF02876; Stap_Strp_tox_C; 1.
DR   PRINTS; PR00279; BACTRLTOXIN.
DR   PROSITE; PS00277; STAPH_STREP_TOXIN_1; 1.
DR   PROSITE; PS00278; STAPH_STREP_TOXIN_2; 1.
SQ   SEQUENCE   271 AA;  31267 MW;  3493F6228B042F10 CRC64;

  Query Match             46.6%;  Score 549;  DB 2;  Length 271;
  Best Local Similarity   47.2%;  Pred. No. 1.3e-32;
  Matches  111;  Conservative   44;  Mismatches   66;  Indels   14;  Gaps    6;

Qy          1 QQDPDPSQLHRSSLVKNL-QNIYFLYEGDPVTHENVKSVDQLLSHHLIYNVSG---PNYD 56
              | || | :||::|     | :|:  ||:   |:   |||||: |:| ||||:|      |||
Db         35 QPDPTPDELHKASKFTGLMENMKVLYDDRYVSATKVKSVDKFLAHDLIYNISDKKLKNYD 94

Qy         57 KLKTELKNQEMATLFKDKNVDIYGVEYYHLCYLC--ENAER----SACIYGGVTNHEGNH 110
              |:|||| |:::|  :||: ||:||  ||  ||     :|  :        |:|||:| |||||
Db         95 KVKTELLNEDLAKKYKDEVVDVYGSNYYVNCYFSSKDNVGKVTGGKTCMYGGITKHEGNH 154

Qy        111 LEIPK--KIVVKVSIDGIQSLSFDIETNKKMVTAQELDYKVRKYLTDNKQLYTNGPSKYE 168
               :  |   ::::|  :   ::||:::|:|| ||||||| | | :| : | ||    | ||
Db        155 FDNGKLQNVLIRVYENKRNTISFEVQTDKKSVTAQELDIKARNFLINKKNLYEFNSSPYE 214

Qy        169 TGYIKFIPKNKESFWFDFFPEP--EFTQSKYLMIYKDNETLDSNTSQIEVYLTTK 221
              |||||||  |  :||:|  | |  :| |||||||:| ||:|:|| : :|||:||||
Db        215 TGYIKFIENNGNTFWYDMMPAPGDKFDQSKYLMMYNDNKTVDSKSVKIEVHLTTK 269


RESULT 11
Q06532
ID   Q06532      PRELIMINARY;      PRT;   239 AA.
AC   Q06532;
DT   01-NOV-1996 (TrEMBLrel. 01, Created)
DT   01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
DT   01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE   Enterotoxin (Fragment).
OS   Staphylococcus aureus.
OC   Bacteria; Firmicutes; Bacillales; Staphylococcus.
OX   NCBI_TaxID=1280;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=740N;
RX   MEDLINE=94011313; PubMed=8406814;
RA   Marr J.C., Lyon J.D., Roberson J.R., Lupher M., Bohach G.A.;
RT   "Characterization of novel type C staphylococcal enterotoxins:
RT   biological and evolutionary implications.";
RL   Infect. Immun. 61:4254-4262(1993).
DR   EMBL; L13375; AAA26619.1; -.
DR   HSSP; P34071; 1SE2.
DR   GO; GO:0005576; C:extracellular; IEA.
DR   GO; GO:0015070; F:toxin activity; IEA.
DR   GO; GO:0009405; P:pathogenesis; IEA.
DR   InterPro; IPR008992; Bact_endotox.
DR   InterPro; IPR006177; Bctrl_tox.
DR   InterPro; IPR006123; Stap/Strep_toxin.
DR   InterPro; IPR006126; Staph/Strept_tox.
DR   InterPro; IPR006173; Staph_tox_OB.
DR   Pfam; PF01123; Stap_Strp_toxin; 1.
DR   Pfam; PF02876; Stap_Strp_tox_C; 1.
DR   PRINTS; PR00279; BACTRLTOXIN.
DR   PROSITE; PS00277; STAPH_STREP_TOXIN_1; 1.
DR   PROSITE; PS00278; STAPH_STREP_TOXIN_2; 1.
FT   NON_TER       1      1
SQ   SEQUENCE   239 AA;  27642 MW;  C77009F46BC8D645 CRC64;

  Query Match             46.5%;  Score 548;  DB 2;  Length 239;
  Best Local Similarity   47.2%;  Pred. No. 1.3e-32;
  Matches  111;  Conservative   43;  Mismatches   67;  Indels   14;  Gaps    6;

Qy          1 QQDPDPSQLHRSS-LVKNLQNIYFLYEGDPVTHENVKSVDQLLSHHLIYNVSG---PNYD 56
              | || |  ||:||     : |: :||:   |:   |||||: |:| ||||:|      |||
Db          3 QPDPMPDDLHKSSEFTGTMDNMKYLYDDHYVSATKVKSVDKFLAHDLIYNISDKKLKNYD 62

Qy         57 KLKTELKNQEMATLFKDKNVDIYGVEYYHLCYLC--ENAER----SACIYGGVTNHEGNH 110
              |:|||| |:::|  :||: ||:||  ||  ||     :|  :        |:|||:| |||||
Db         63 KVKTELLNEDLAKKYKDEVVDVYGSNYYVNCYFSSKDNVGKVTGGKTCMYGGITKHEGNH 122

Qy        111 LEIP--KKIVVKVSIDGIQSLSFDIETNKKMVTAQELDYKVRKYLTDNKQLYTNGPSKYE 168
               :    : ::::|  :   ::||:::|:|| ||||||| | | :| : | ||    | ||
Db        123 FDNGNLQNVLIRVYENKRNTISFEVQTDKKSVTAQELDIKARNFLINKKNLYEFNSSPYE 182

Qy        169 TGYIKFIPKNKESFWFDFFPEP--EFTQSKYLMIYKDNETLDSNTSQIEVYLTTK 221
              |||||||  |  :||:|  | |  :| |||||||:| ||:|:||   :|||:||||
Db        183 TGYIKFIENNGNTFWYDMMPAPGDKFDQSKYLMMYNDNKTVDSKRVKIEVHLTTK 237


RESULT 12
Q06533
ID   Q06533      PRELIMINARY;      PRT;   239 AA.
AC   Q06533;
DT   01-NOV-1996 (TrEMBLrel. 01, Created)
DT   01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
DT   01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE   Enterotoxin (Fragment).
OS   Staphylococcus aureus.
OC   Bacteria; Firmicutes; Bacillales; Staphylococcus.
OX   NCBI_TaxID=1280;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=MNCopeland;
RX   MEDLINE=94011313; PubMed=8406814;
RA   Marr J.C., Lyon J.D., Roberson J.R., Lupher M., Bohach G.A.;
RT   "Characterization of novel type C staphylococcal enterotoxins:
RT   biological and evolutionary implications.";
RL   Infect. Immun. 61:4254-4262(1993).
DR   EMBL; L13378; AAA26622.1; -.
DR   HSSP; P34071; 1STE.
DR   GO; GO:0005576; C:extracellular; IEA.
DR   GO; GO:0015070; F:toxin activity; IEA.
DR   GO; GO:0009405; P:pathogenesis; IEA.
DR   InterPro; IPR008992; Bact_endotox.
DR   InterPro; IPR006177; Bctrl_tox.
DR   InterPro; IPR006123; Stap/Strep_toxin.
DR   InterPro; IPR006126; Staph/Strept_tox.
DR   InterPro; IPR006173; Staph_tox_OB.
DR   Pfam; PF01123; Stap_Strp_toxin; 1.
DR   Pfam; PF02876; Stap_Strp_tox_C; 1.
DR   PRINTS; PR00279; BACTRLTOXIN.
DR   PROSITE; PS00277; STAPH_STREP_TOXIN_1; 1.
DR   PROSITE; PS00278; STAPH_STREP_TOXIN_2; 1.
FT   NON_TER       1      1
SQ   SEQUENCE   239 AA;  27651 MW;  A21A954386AE8625 CRC64;

  Query Match             46.4%;  Score 547;  DB 2;  Length 239;
  Best Local Similarity   46.8%;  Pred. No. 1.6e-32;
  Matches  110;  Conservative   44;  Mismatches   67;  Indels   14;  Gaps    6;

Qy          1 QQDPDPSQLHRSS-LVKNLQNIYFLYEGDPVTHENVKSVDQLLSHHLIYNVSG---PNYD 56
              | || | :||:||     : |: :||:   |:   |||||: |:| ||||:|      |||
Db          3 QPDPTPDELHKSSEFTGTMGNMKYLYDDHYVSATKVKSVDKFLAHDLIYNISDKKLKNYD 62

Qy         57 KLKTELKNQEMATLFKDKNVDIYGVEYYHLCYLC--ENAER----SACIYGGVTNHEGNH 110
              |:|||| |:::|  :||: ||:||  ||  ||     :|  :        |:|||:| |||||
Db         63 KVKTELLNEDLAKKYKDEVVDVYGSNYYVNCYFSSKDNVGKVTGGKTCMYGGITKHEGNH 122

Qy        111 LEIP--KKIVVKVSIDGIQSLSFDIETNKKMVTAQELDYKVRKYLTDNKQLYTNGPSKYE 168
               :    : ::::|  :   ::||:::|:|| ||||||| | | :| : | ||    | ||
Db        123 FDNGNLQNVLIRVYENKRNTISFEVQTDKKSVTAQELDIKARNFLINKKNLYEFNSSPYE 182

Qy        169 TGYIKFIPKNKESFWFDFFPEP--EFTQSKYLMIYKDNETLDSNTSQIEVYLTTK 221
              |||||||  |  :||:|  | |  :| |||||||:| ||:|:||   :|||:||||
Db        183 TGYIKFIENNGNTFWYDMMPAPGDKFDQSKYLMMYNDNKTVDSKRVKIEVHLTTK 237
```

```
RESULT 13
Q8NXJ6
ID   Q8NXJ6      PRELIMINARY;      PRT;   266 AA.
AC   Q8NXJ6;
DT   01-OCT-2002 (TrEMBLrel. 22, Created)
DT   01-OCT-2002 (TrEMBLrel. 22, Last sequence update)
DT   01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE   Enterotoxin type C precursor.
GN   SEC4 OR MW0759.
OS   Staphylococcus aureus (strain MW2).
OC   Bacteria; Firmicutes; Bacillales; Staphylococcus.
OX   NCBI_TaxID=196620;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=22040717; PubMed=12044378;
RA   Baba T., Takeuchi F., Kuroda M., Yuzawa H., Aoki K.-I., Oguchi A.,
RA   Nagai Y., Iwama N., Asano K., Naimi T., Kuroda H., Cui L.,
RA   Yamamoto K., Hiramatsu K.;
RT   "Genome and virulence determinants of high virulence community-
RT   acquired MRSA.";
RL   Lancet 359:1819-1827(2002).
DR   EMBL; AP004824; BAB94624.1; -.
DR   GO; GO:0005576; C:extracellular; IEA.
DR   GO; GO:0015070; F:toxin activity; IEA.
DR   GO; GO:0009405; P:pathogenesis; IEA.
DR   InterPro; IPR008992; Bact_endotox.
DR   InterPro; IPR006177; Bctrl_tox.
DR   InterPro; IPR006123; Stap/Strep_toxin.
DR   InterPro; IPR006126; Staph/Strept_tox.
DR   InterPro; IPR006173; Staph_tox_OB.
DR   Pfam; PF01123; Stap_Strp_toxin; 1.
DR   Pfam; PF02876; Stap_Strp_tox_C; 1.
DR   PRINTS; PR00279; BACTRLTOXIN.
DR   PROSITE; PS00277; STAPH_STREP_TOXIN_1; 1.
DR   PROSITE; PS00278; STAPH_STREP_TOXIN_2; 1.
KW   Complete proteome.
SQ   SEQUENCE   266 AA;  30670 MW;  4C654659AA48120F CRC64;

  Query Match             46.4%;  Score 547;  DB 16;  Length 266;
  Best Local Similarity   46.8%;  Pred. No. 1.8e-32;
  Matches  110;  Conservative   44;  Mismatches   67;  Indels   14;  Gaps    6;

Qy          1 QQDPDPSQLHRSS-LVKNLQNIYFLYEGDPVTHENVKSVDQLLSHHLIYNVSG---PNYD 56
              | || | :||:||     : |: :||:   |:   |||||: |:| ||||:|     |||
Db         30 QPDPTPDELHKSSEFTGTMGNMKYLYDDHYVSATKVKSVDKFLAHDLIYNISDKKLKNYD 89

Qy         57 KLKTELKNQEMATLFKDKNVDIYGVEYYHLCYLC--ENAER----SACIYGGVTNHEGNH 110
              |:|||| |:::|  :||: ||:||  ||  ||    :|  :       |:|||:| |||||
Db         90 KVKTELLNEDLAKKYKDEVVDVYGSNYYVNCYFSSKDNVGKVTGGKTCMYGGITKHEGNH 149

Qy        111 LEIP--KKIVVKVSIDGIQSLSFDIETNKKMVTAQELDYKVRKYLTDNKQLYTNGPSKYE 168
                   :    : ::::|  :   ::||::|:|| ||||||| | | :| : | ||    | ||
Db        150 FDNGNLQNVLIRVYENKRNTISFEVQTDKKSVTAQELDIKARNFLINKKNLYEFNSSPYE 209

Qy        169 TGYIKFIPKNKESFWFDFFPEP--EFTQSKYLMIYKDNETLDSNTSQIEVYLTTK 221
              |||||||  |  :||:|  | |  :| |||||||:| ||:|:||  :|||:||||
Db        210 TGYIKFIENNGNTFWYDMMPAPGDKFDQSKYLMMYNDNKTVDSKRVKIEVHLTTK 264


RESULT 14
O05157
ID   O05157      PRELIMINARY;      PRT;   239 AA.
AC   O05157;
DT   01-JUL-1997 (TrEMBLrel. 04, Created)
DT   01-JUL-1997 (TrEMBLrel. 04, Last sequence update)
DT   01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE   Type C enterotoxin (Fragment).
OS   Staphylococcus intermedius.
OC   Bacteria; Firmicutes; Bacillales; Staphylococcus.
OX   NCBI_TaxID=1285;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=95-011195;
RA   Edwards V.M., Deringer J.R., Callantine S.D., Deobald C.F.,
RA   Berger P.H., Kapur V., Stauffacher C.V., Bohach G.A.;
RT   "Characterization of the type C enterotoxin (SEC-canine) produced by
RT   Staphylococcus intermedius pyoderma isolates.";
RL   Submitted (FEB-1997) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; U91526; AAB50248.1; -.
DR   HSSP; P23313; 1JCK.
DR   GO; GO:0005576; C:extracellular; IEA.
DR   GO; GO:0015070; F:toxin activity; IEA.
DR   GO; GO:0009405; P:pathogenesis; IEA.
DR   InterPro; IPR008992; Bact_endotox.
DR   InterPro; IPR006177; Bctrl_tox.
DR   InterPro; IPR006123; Stap/Strep_toxin.
DR   InterPro; IPR006126; Staph/Strept_tox.
DR   InterPro; IPR006173; Staph_tox_OB.
DR   Pfam; PF01123; Stap_Strp_toxin; 1.
DR   Pfam; PF02876; Stap_Strp_tox_C; 1.
DR   PRINTS; PR00279; BACTRLTOXIN.
DR   PROSITE; PS00277; STAPH_STREP_TOXIN_1; 1.
DR   PROSITE; PS00278; STAPH_STREP_TOXIN_2; 1.
FT   NON_TER       1       1
SQ   SEQUENCE   239 AA;  27536 MW;  D660644660DE4191 CRC64;

  Query Match             46.3%;  Score 546;  DB 2;  Length 239;
  Best Local Similarity   47.7%;  Pred. No. 1.9e-32;
  Matches  112;  Conservative   41;  Mismatches   68;  Indels   14;  Gaps    6;

Qy          1 QQDPDPSQLHRSS-LVKNLQNIYFLYEGDPVTHENVKSVDQLLSHHLIYNVSG---PNYD 56
              | || | :||:||     : |: :||:   |:   |||||: |:| ||||:|     |||
Db          3 QPDPMPDDLHKSSEFTGTMGNMKCLYDDYYVSATKVKSVDKFLAHDLIYNISDKKLKNYD 62

Qy         57 KLKTELKNQEMATLFKDKNVDIYGVEYYHLCYLC--ENAER----SACIYGGVTNHEGNH 110
              |:|||| |:::|  :||: ||:||  ||  ||    :|  :       |:|||:| |||||
Db         63 KVKTELLNEDLAKKYKDEVVDVYGSNYYVNCYFSSKDNVGKVTGGKTCMYGGITKHEGNH 122

Qy        111 LEIP--KKIVVKVSIDGIQSLSFDIETNKKMVTAQELDYKVRKYLTDNKQLYTNGPSKYE 168
                   :    : ::::|  :   ::||::|:|| ||||||| | | :| : | ||    | ||
Db        123 FDNGNLQNVLIRVYENKRNTISFDVQTDKKSVTAQELDIKARNFLINKKNLYEFNSSPYE 182

Qy        169 TGYIKFIPKNKESFWFDFFPEP--EFTQSKYLMIYKDNETLDSNTSQIEVYLTTK 221
              |||||||  |  :||:|  | |  :| |||||||:| ||:|:||  :|||:||||
Db        183 TGYIKFIESNDNTFWYDMMPAPGDKFDQSKYLMIYSDNKTVDSKSVKIEVHLTTK 237


RESULT 15
Q06531
ID   Q06531      PRELIMINARY;      PRT;   239 AA.
AC   Q06531;
DT   01-NOV-1996 (TrEMBLrel. 01, Created)
DT   01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
DT   01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE   Enterotoxin (Fragment).
OS   Staphylococcus aureus.
OC   Bacteria; Firmicutes; Bacillales; Staphylococcus.
OX   NCBI_TaxID=1280;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=4446;
RX   MEDLINE=94011313; PubMed=8406814;
RA   Marr J.C., Lyon J.D., Roberson J.R., Lupher M., Bohach G.A.;
RT   "Characterization of novel type C staphylococcal enterotoxins:
RT   biological and evolutionary implications.";
RL   Infect. Immun. 61:4254-4262(1993).
DR   EMBL; L13374; AAA26618.1; -.
DR   HSSP; P34071; 1SE2.
DR   GO; GO:0005576; C:extracellular; IEA.
DR   GO; GO:0015070; F:toxin activity; IEA.
DR   GO; GO:0009405; P:pathogenesis; IEA.
DR   InterPro; IPR008992; Bact_endotox.
DR   InterPro; IPR006177; Bctrl_tox.
DR   InterPro; IPR006123; Stap/Strep_toxin.
DR   InterPro; IPR006126; Staph/Strept_tox.
DR   InterPro; IPR006173; Staph_tox_OB.
```

```
DR   Pfam; PF01123; Stap_Strp_toxin; 1.
DR   Pfam; PF02876; Stap_Strp_tox_C; 1.
DR   PRINTS; PR00279; BACTRLTOXIN.
DR   PROSITE; PS00277; STAPH_STREP_TOXIN_1; 1.
DR   PROSITE; PS00278; STAPH_STREP_TOXIN_2; 1.
FT   NON_TER       1      1
SQ   SEQUENCE   239 AA;  27612 MW;  BB7BD6204731ED24 CRC64;

  Query Match             46.3%;  Score 546;  DB 2;  Length 239;
  Best Local Similarity   47.2%;  Pred. No. 1.9e-32;
  Matches  111;  Conservative   43;  Mismatches   67;  Indels   14;  Gaps    6;

Qy          1 QQDPDPSQLHRSS-LVKNLQNIYFLYEGDPVTHENVKSVDQLLSHHLIYNVSG---PNYD 56
              | || |  ||:||     : |: :||:   |:   |||||: |:| ||||:|     |||
Db          3 QPDPMPDDLHKSSEFTGTMGNMKYLYDDHYVSATKVKSVDKFLAHDLIYNISDKRLKNYD 62

Qy         57 KLKTELKNQEMATLFKDKNVDIYGVEYYHLCYLC--ENAER----SACIYGGVTNHEGNH 110
              |:|||| |:::|  :||: ||:||  ||  ||    :|  :      |:|||:| |||||
Db         63 KVKTELLNEDLAKKYKDEVVDVYGSNYYVNCYFSSKDNVGKVTGGKTCMYGGITKHEGNH 122

Qy        111 LEIP--KKIVVKVSIDGIQSLSFDIETNKKMVTAQELDYKVRKYLTDNKQLYTNGPSKYE 168
              :    : ::|:|  :   ::||:::|:|| ||||||| | | :| : | ||    | ||
Db        123 FDNGNLQNVLVRVYENKRNTISFEVQTDKKSVTAQELDIKARNFLINKKNLYEFNSSPYE 182

Qy        169 TGYIKFIPKNKESFWFDFFPEP--EFTQSKYLMIYKDNETLDSNTSQIEVYLTTK 221
              |||||||  |  :||:|  | |  :| ||||||:| ||:|:|| : :|||:||||
Db        183 TGYIKFIENNGNTFWYDMMPAPGDKFDQSKYLMMYNDNKTVDSKSVKIEVHLTTK 237


Search completed: May  7, 2004, 12:09:08
Job time : 49 secs
```

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: May 7, 2004, 12:00:52 ; Search time 60 Seconds
(without alignments)
1040.717 Million cell updates/sec

Title: US-10-625-221-14
Perfect score: 1178
Sequence: 1 QQDPDPSQLHRSSLVKNLQN.....KDNETLDSNTSQIEVYLTTK 221

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 1586107 seqs, 282547505 residues

Total number of hits satisfying chosen parameters: 1586107

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : A_Geneseq_29Jan04:*
1: geneseqp1980s:*
2: geneseqp1990s:*
3: geneseqp2000s:*
4: geneseqp2001s:*
5: geneseqp2002s:*
6: geneseqp2003as:*
7: geneseqp2003bs:*
8: geneseqp2004s:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	1178	100.0	251	2	AAW12097	Aaw12097 Streptoco
2	1178	100.0	251	2	AAW59780	Aaw59780 Amino aci
3	1175	99.7	251	2	AAW12154	Aaw12154 Streptoco
4	1174	99.7	251	2	AAW12146	Aaw12146 Streptoco
5	1173	99.6	251	2	AAW12150	Aaw12150 Streptoco
6	1173	99.6	251	2	AAW12147	Aaw12147 Streptoco
7	1169	99.2	221	4	AAB67344	Aab67344 Streptoco
8	1169	99.2	251	2	AAW12148	Aaw12148 Streptoco
9	1168	99.2	251	2	AAW12153	Aaw12153 Streptoco
10	1168	99.2	251	2	AAW12151	Aaw12151 Streptoco
11	1168	99.2	251	2	AAW12152	Aaw12152 Streptoco
12	1164	98.8	251	7	ABU62460	Abu62460 Streptoco
13	1163	98.7	221	2	AAR13209	Aar13209 Streptoco
14	1163	98.7	221	2	AAR45017	Aar45017 Staphyloc
15	1163	98.7	221	5	ABB76240	Abb76240 Staphyloc
16	1163	98.7	251	2	AAW12149	Aaw12149 Streptoco
17	1163	98.7	251	2	AAW59798	Aaw59798 Amino aci
18	1163	98.7	251	2	AAW59781	Aaw59781 Amino aci
19	1163	98.7	251	3	AAY70109	Aay70109 Streptoco
20	1163	98.7	251	5	ABB79508	Abb79508 Streptoco
21	1163	98.7	251	6	ABU10088	Abu10088 Streptoco
22	1163	98.7	251	7	ABU62331	Abu62331 Streptoco
23	1163	98.7	251	7	AAE37683	Aae37683 Streptoco
24	1158.5	98.3	250	2	AAW12145	Aaw12145 Streptoco
25	1122	95.2	220	7	ABU62334	Abu62334 Streptoco
26	1122	95.2	220	7	AAE37687	Aae37687 Streptoco
27	1122	95.2	468	7	ABU62335	Abu62335 SPEa L42R
28	1122	95.2	468	7	AAE37684	Aae37684 Streptoco
29	1117	94.8	220	7	AAE37689	Aae37689 S. pyogen
30	1116	94.7	220	7	AAE37688	Aae37688 S. pyogen
31	1116	94.7	468	7	AAE37691	Aae37691 S. pyogen
32	978.5	83.1	250	6	ABU79074	Abu79074 S. pyogen
33	574	48.7	239	2	AAW64647	Aaw64647 Synthetic
34	572	48.6	239	4	AAB67341	Aab67341 Staphyloc
35	572	48.6	255	2	AAW06737	Aaw06737 Staphyloc
36	572	48.6	266	6	ABU79069	Abu79069 S. aureus
37	570	48.4	266	3	AAY92319	Aay92319 Plant-opt
38	563	47.8	266	5	ABB79503	Abb79503 Staphyloc
39	563	47.8	266	6	ABU10083	Abu10083 Staphyloc
40	563	47.8	266	7	ABU62326	Abu62326 S. aureus
41	563	47.8	266	7	AAE37678	Aae37678 Protein #
42	560	47.5	239	5	ABB76237	Abb76237 Staphyloc
43	560	47.5	266	7	ABU62453	Abu62453 S. aureus
44	559	47.5	239	2	AAR13206	Aar13206 Staphyloc
45	559	47.5	239	2	AAR45014	Aar45014 Staphyloc

ALIGNMENTS

RESULT 1
AAW12097
ID AAW12097 standard; protein; 251 AA.
XX
AC AAW12097;
XX
DT 04-NOV-1997 (first entry)
XX
DE Streptococcus pyogenes Streptococcal toxin A.
XX
KW Streptococcal; toxin A; SPE-A; non-lethal; mutant; production; vaccine;
KW protection; treatment; cancer; neutralising antibody;
KW streptococcal toxic shock syndrome; STSS; symptom; amelioration; fever;
KW hypotension; group A streptococcal infection; myositis; fascitis;
KW liver damage; T cell; lymphoma; ovarian; uterine.
XX
OS Streptococcus pyogenes.
XX
FH Key Location/Qualifiers
FT Peptide 1. .30
FT /label= sig_peptide
FT Peptide 31. .251
FT /label= mat_peptide
XX
PN WO9640930-A1.
XX
PD 19-DEC-1996.
XX
PF 07-JUN-1996; 96WO-US010252.
XX
PR 07-JUN-1995; 95US-00480261.
XX
PA (MINU) UNIV MINNESOTA.
XX
PI Schlievert PM, Roggiani M, Stoehr J, Ohlendorf D;
XX
DR WPI; 1997-099936/09.
DR N-PSDB; AAW12097.
XX
PT Mutant SPE-A toxin with at least one amino acid change is substantially
PT non-lethal - used in vaccine composition for treatment of cancer and
PT streptococcal toxic shock syndrome etc.
XX
PS Disclosure; Page 77-79; 102pp; English.
XX
CC The present sequence is Streptococcus pyogenes Streptococcal toxin A (SPE
CC -A), from which a non-lethal mutant SPE-A, comprising at least 1 amino
CC acid change, can be derived. The mutant SPE-A can be used to produce

```
CC   vaccines to protect animals against wild type SPE-A and to treat cancer
CC   and streptococcal toxic shock syndrome (STSS). The mutant SPE-A causes
CC   neutralising antibodies (Ab) to be produced, which may be used to
CC   ameliorate STSS symptoms, e.g. fever, hypotension, group A streptococcal
CC   infection, myositis, fascitis and liver damage. The neutralising Ab is
CC   preferably administered in conjunction with antibiotic therapy. The
CC   mutant SPE-A is especially useful for treating T cell lymphomas, and
CC   ovarian and uterine cancer. It is thought that mutant SPE-A can be
CC   selectively toxic to T cell lymphoma cells
XX
SQ   Sequence 251 AA;

  Query Match             100.0%;  Score 1178;  DB 2;  Length 251;
  Best Local Similarity   100.0%;  Pred. No. 8.3e-101;
  Matches  221;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 QQDPDPSQLHRSSLVKNLQNIYFLYEGDPVTHENVKSVDQLLSHHLIYNVSGPNYDKLKT 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         31 QQDPDPSQLHRSSLVKNLQNIYFLYEGDPVTHENVKSVDQLLSHHLIYNVSGPNYDKLKT 90

Qy         61 ELKNQEMATLFKDKNVDIYGVEYYHLCYLCENAERSACIYGGVTNHEGNHLEIPKKIVVK 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         91 ELKNQEMATLFKDKNVDIYGVEYYHLCYLCENAERSACIYGGVTNHEGNHLEIPKKIVVK 150

Qy        121 VSIDGIQSLSFDIETNKKMVTAQELDYKVRKYLTDNKQLYTNGPSKYETGYIKFIPKNKE 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        151 VSIDGIQSLSFDIETNKKMVTAQELDYKVRKYLTDNKQLYTNGPSKYETGYIKFIPKNKE 210

Qy        181 SFWFDFFPEPEFTQSKYLMIYKDNETLDSNTSQIEVYLTTK 221
              |||||||||||||||||||||||||||||||||||||||||
Db        211 SFWFDFFPEPEFTQSKYLMIYKDNETLDSNTSQIEVYLTTK 251


RESULT 2
AAW59780
ID   AAW59780 standard; protein; 251 AA.
XX
AC   AAW59780;
XX
DT   12-OCT-1998  (first entry)
XX
DE   Amino acid sequence of Streptococcus pyogenes exotoxin A.
XX
KW   SPE-A toxin; nonlethal; mutant; Streptococcus pyogenes exotoxin A;
KW   wild type; nontoxic; antibody; vaccine; immunity; ovarian cancer;
KW   streptococcal toxic shock syndrome; STSS; T cell lymphoma;
KW   uterine cancer.
XX
OS   Streptococcus pyogenes.
XX
PN   WO9824911-A2.
XX
PD   11-JUN-1998.
XX
PF   05-DEC-1997;   97WO-US022228.
XX
PR   06-DEC-1996;   96US-0032930P.
XX
PA   (MINU ) UNIV MINNESOTA.
XX
PI   Schlievert PM,  Roggiani M,  Stoehr J,  Ohlendorf D;
XX
DR   WPI; 1998-333330/29.
DR   N-PSDB; AAV41593.
XX
PT   New mutant Streptococcal SPE-A toxins - useful for, e.g. prevention or
PT   treatment of streptococcal infection or toxic shock syndrome.
XX
PS   Disclosure; Fig 3; 95pp; English.
XX
CC   This is the amino acid sequence of the Streptococcus pyogenes exotoxin A
CC   (SPE-A toxin). The novel mutant Streptococcal SPE-A toxin has at least 1
CC   aa change and is nonlethal compared with a protein to wild type SPE-A
CC   toxin. The mutant SPE-A toxins are nontoxic and can produce antibodies
CC   that neutralise wild type SPE-A toxin activity. The toxins can be used in
CC   vaccines and therapeutics to generate a protective immune response
CC   against streptococcal infection. They can be used to protect against the
CC   development of streptococcal toxic shock syndrome (STSS). In addition,
CC   the toxins can be used for treating animals with symptoms of
CC   streptococcal infection or STSS and in methods for stimulating T cell
CC   proliferation and in the treatment of cancer. In particular they can be
CC   used for treating T cell lymphomas, and ovarian and uterine cancer
XX
SQ   Sequence 251 AA;

  Query Match             100.0%;  Score 1178;  DB 2;  Length 251;
  Best Local Similarity   100.0%;  Pred. No. 8.3e-101;
  Matches  221;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 QQDPDPSQLHRSSLVKNLQNIYFLYEGDPVTHENVKSVDQLLSHHLIYNVSGPNYDKLKT 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         31 QQDPDPSQLHRSSLVKNLQNIYFLYEGDPVTHENVKSVDQLLSHHLIYNVSGPNYDKLKT 90

Qy         61 ELKNQEMATLFKDKNVDIYGVEYYHLCYLCENAERSACIYGGVTNHEGNHLEIPKKIVVK 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         91 ELKNQEMATLFKDKNVDIYGVEYYHLCYLCENAERSACIYGGVTNHEGNHLEIPKKIVVK 150

Qy        121 VSIDGIQSLSFDIETNKKMVTAQELDYKVRKYLTDNKQLYTNGPSKYETGYIKFIPKNKE 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        151 VSIDGIQSLSFDIETNKKMVTAQELDYKVRKYLTDNKQLYTNGPSKYETGYIKFIPKNKE 210

Qy        181 SFWFDFFPEPEFTQSKYLMIYKDNETLDSNTSQIEVYLTTK 221
              |||||||||||||||||||||||||||||||||||||||||
Db        211 SFWFDFFPEPEFTQSKYLMIYKDNETLDSNTSQIEVYLTTK 251


RESULT 3
AAW12154
ID   AAW12154 standard; protein; 251 AA.
XX
AC   AAW12154;
XX
DT   04-NOV-1997  (first entry)
XX
DE   Streptococcus pyogenes Streptococcal toxin A mutant Ser195Ala.
XX
KW   Streptococcal; toxin A; SPE-A; non-lethal; mutant; production; vaccine;
KW   protection; treatment; cancer; neutralising antibody;
KW   streptococcal toxic shock syndrome; STSS; symptom; amelioration; fever;
KW   hypotension; group A streptococcal infection; myositis; fascitis;
KW   liver damage; T cell; lymphoma; ovarian; uterine.
XX
OS   Streptococcus pyogenes.
OS   Synthetic.
XX
FH   Key             Location/Qualifiers
FT   Peptide         1. .30
FT                   /label= sig_peptide
FT   Peptide         31. .251
FT                   /label= mat_peptide
FT   Misc-difference 225
FT                   /note= "wild type Ser replaced by Ala"
XX
PN   WO9640930-A1.
XX
PD   19-DEC-1996.
XX
PF   07-JUN-1996;   96WO-US010252.
XX
PR   07-JUN-1995;   95US-00480261.
XX
PA   (MINU ) UNIV MINNESOTA.
XX
PI   Schlievert PM,  Roggiani M,  Stoehr J,  Ohlendorf D;
XX
DR   WPI; 1997-099936/09.
```

XX
PT Mutant SPE-A toxin with at least one amino acid change is substantially
PT non-lethal - used in vaccine composition for treatment of cancer and
PT streptococcal toxic shock syndrome etc.
XX
PS Example 4; Page; 102pp; English.
XX
CC The present sequence is a non-lethal Streptococcus pyogenes Streptococcal
CC toxin A (SPE-A) mutant, which can be used to produce vaccines to protect
CC animals against wild type SPE-A and to treat cancer and streptococcal
CC toxic shock syndrome (STSS). The mutant SPE-A causes neutralising
CC antibodies (Ab) to be produced, which may be used to ameliorate STSS
CC symptoms, e.g. fever, hypotension, group A streptococcal infection,
CC myositis, fascitis and liver damage. The neutralising Ab is preferably
CC administered in conjunction with antibiotic therapy. The mutant SPE-A is
CC especially useful for treating T cell lymphomas, and ovarian and uterine
CC cancer. It is thought that mutant SPE-A can be selectively toxic to T
CC cell lymphoma cells. N.B. Sequence not given in the specification, but
CC constructed using the wild type SPE-A sequence given on pages 77-79
XX
SQ Sequence 251 AA;

Query Match 99.7%; Score 1175; DB 2; Length 251;
Best Local Similarity 99.5%; Pred. No. 1.6e-100;
Matches 220; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

```
Qy          1 QQDPDPSQLHRSSLVKNLQNIYFLYEGDPVTHENVKSVDQLLSHHLIYNVSGPNYDKLKT 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         31 QQDPDPSQLHRSSLVKNLQNIYFLYEGDPVTHENVKSVDQLLSHHLIYNVSGPNYDKLKT 90

Qy         61 ELKNQEMATLFKDKNVDIYGVEYYHLCYLCENAERSACIYGGVTNHEGNHLEIPKKIVVK 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         91 ELKNQEMATLFKDKNVDIYGVEYYHLCYLCENAERSACIYGGVTNHEGNHLEIPKKIVVK 150

Qy        121 VSIDGIQSLSFDIETNKKMVTAQELDYKVRKYLTDNKQLYTNGPSKYETGYIKFIPKNKE 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        151 VSIDGIQSLSFDIETNKKMVTAQELDYKVRKYLTDNKQLYTNGPSKYETGYIKFIPKNKE 210

Qy        181 SFWFDFFPEPEFTQSKYLMIYKDNETLDSNTSQIEVYLTTK 221
              |||||||||||||:|||||||||||||||||||||||||||
Db        211 SFWFDFFPEPEFTQAKYLMIYKDNETLDSNTSQIEVYLTTK 251
```

RESULT 4
AAW12146
ID AAW12146 standard; protein; 251 AA.
XX
AC AAW12146;
XX
DT 04-NOV-1997 (first entry)
XX
DE Streptococcus pyogenes Streptococcal toxin A mutant Lys157Glu.
XX
KW Streptococcal; toxin A; SPE-A; non-lethal; mutant; production; vaccine;
KW protection; treatment; cancer; neutralising antibody;
KW streptococcal toxic shock syndrome; STSS; symptom; amelioration; fever;
KW hypotension; group A streptococcal infection; myositis; fascitis;
KW liver damage; T cell; lymphoma; ovarian; uterine.
XX
OS Streptococcus pyogenes.
OS Synthetic.
XX
FH Key Location/Qualifiers
FT Peptide 1. .30
FT /label= sig_peptide
FT Peptide 31. .251
FT /label= mat_peptide
FT Misc-difference 187
FT /note= "wild type Lys replaced by Glu"
XX
PN WO9640930-A1.
XX
PD 19-DEC-1996.
XX
PF 07-JUN-1996; 96WO-US010252.
XX
PR 07-JUN-1995; 95US-00480261.
XX
PA (MINU) UNIV MINNESOTA.
XX
PI Schlievert PM, Roggiani M, Stoehr J, Ohlendorf D;
XX
DR WPI; 1997-099936/09.
XX
PT Mutant SPE-A toxin with at least one amino acid change is substantially
PT non-lethal - used in vaccine composition for treatment of cancer and
PT streptococcal toxic shock syndrome etc.
XX
PS Claim 5; Page; 102pp; English.
XX
CC The present sequence is a non-lethal Streptococcus pyogenes Streptococcal
CC toxin A (SPE-A) mutant, which can be used to produce vaccines to protect
CC animals against wild type SPE-A and to treat cancer and streptococcal
CC toxic shock syndrome (STSS). The mutant SPE-A causes neutralising
CC antibodies (Ab) to be produced, which may be used to ameliorate STSS
CC symptoms, e.g. fever, hypotension, group A streptococcal infection,
CC myositis, fascitis and liver damage. The neutralising Ab is preferably
CC administered in conjunction with antibiotic therapy. The mutant SPE-A is
CC especially useful for treating T cell lymphomas, and ovarian and uterine
CC cancer. It is thought that mutant SPE-A can be selectively toxic to T
CC cell lymphoma cells. N.B. Sequence not given in the specification, but
CC constructed using the wild type SPE-A sequence given on pages 77-79
XX
SQ Sequence 251 AA;

Query Match 99.7%; Score 1174; DB 2; Length 251;
Best Local Similarity 99.5%; Pred. No. 1.9e-100;
Matches 220; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

```
Qy          1 QQDPDPSQLHRSSLVKNLQNIYFLYEGDPVTHENVKSVDQLLSHHLIYNVSGPNYDKLKT 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         31 QQDPDPSQLHRSSLVKNLQNIYFLYEGDPVTHENVKSVDQLLSHHLIYNVSGPNYDKLKT 90

Qy         61 ELKNQEMATLFKDKNVDIYGVEYYHLCYLCENAERSACIYGGVTNHEGNHLEIPKKIVVK 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         91 ELKNQEMATLFKDKNVDIYGVEYYHLCYLCENAERSACIYGGVTNHEGNHLEIPKKIVVK 150

Qy        121 VSIDGIQSLSFDIETNKKMVTAQELDYKVRKYLTDNKQLYTNGPSKYETGYIKFIPKNKE 180
              ||||||||||||||||||||||||||||||||||||:|||||||||||||||||||||||
Db        151 VSIDGIQSLSFDIETNKKMVTAQELDYKVRKYLTDNEQLYTNGPSKYETGYIKFIPKNKE 210

Qy        181 SFWFDFFPEPEFTQSKYLMIYKDNETLDSNTSQIEVYLTTK 221
              |||||||||||||||||||||||||||||||||||||||||||
Db        211 SFWFDFFPEPEFTQSKYLMIYKDNETLDSNTSQIEVYLTTK 251
```

RESULT 5
AAW12150
ID AAW12150 standard; protein; 251 AA.
XX
AC AAW12150;
XX
DT 04-NOV-1997 (first entry)
XX
DE Streptococcus pyogenes Streptococcal toxin A mutant Lys16Asn.
XX
KW Streptococcal; toxin A; SPE-A; non-lethal; mutant; production; vaccine;
KW protection; treatment; cancer; neutralising antibody;
KW streptococcal toxic shock syndrome; STSS; symptom; amelioration; fever;
KW hypotension; group A streptococcal infection; myositis; fascitis;
KW liver damage; T cell; lymphoma; ovarian; uterine.
XX
OS Streptococcus pyogenes.
OS Synthetic.
XX
FH Key Location/Qualifiers

```
FT   Peptide          1. .30
FT                    /label= sig_peptide
FT   Peptide          31. .251
FT                    /label= mat_peptide
FT   Misc-difference 46
FT                    /note= "wild type Lys replaced by Asn"
XX
PN   WO9640930-A1.
XX
PD   19-DEC-1996.
XX
PF   07-JUN-1996;   96WO-US010252.
XX
PR   07-JUN-1995;   95US-00480261.
XX
PA   (MINU ) UNIV MINNESOTA.
XX
PI   Schlievert PM,  Roggiani M,  Stoehr J,  Ohlendorf D;
XX
DR   WPI; 1997-099936/09.
XX
PT   Mutant SPE-A toxin with at least one amino acid change is substantially
PT   non-lethal - used in vaccine composition for treatment of cancer and
PT   streptococcal toxic shock syndrome etc.
XX
PS   Example 4; Page; 102pp; English.
XX
CC   The present sequence is a non-lethal Streptococcus pyogenes Streptococcal
CC   toxin A (SPE-A) mutant, which can be used to produce vaccines to protect
CC   animals against wild type SPE-A and to treat cancer and streptococcal
CC   toxic shock syndrome (STSS). The mutant SPE-A causes neutralising
CC   antibodies (Ab) to be produced, which may be used to ameliorate STSS
CC   symptoms, e.g. fever, hypotension, group A streptococcal infection,
CC   myositis, fascitis and liver damage. The neutralising Ab is preferably
CC   administered in conjunction with antibiotic therapy. The mutant SPE-A is
CC   especially useful for treating T cell lymphomas, and ovarian and uterine
CC   cancer. It is thought that mutant SPE-A can be selectively toxic to T
CC   cell lymphoma cells. N.B. Sequence not given in the specification, but
CC   constructed using the wild type SPE-A sequence given on pages 77-79
XX
SQ   Sequence 251 AA;

  Query Match             99.6%;  Score 1173;  DB 2;  Length 251;
  Best Local Similarity   99.5%;  Pred. No. 2.4e-100;
  Matches  220;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy          1 QQDPDPSQLHRSSLVKNLQNIYFLYEGDPVTHENVKSVDQLLSHHLIYNVSGPNYDKLKT 60
              ||||||||||||||| |||||||||||||||||||||||||||||||||||||||||||||
Db         31 QQDPDPSQLHRSSLVNNLQNIYFLYEGDPVTHENVKSVDQLLSHHLIYNVSGPNYDKLKT 90

Qy         61 ELKNQEMATLFKDKNVDIYGVEYYHLCYLCENAERSACIYGGVTNHEGNHLEIPKKIVVK 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         91 ELKNQEMATLFKDKNVDIYGVEYYHLCYLCENAERSACIYGGVTNHEGNHLEIPKKIVVK 150

Qy        121 VSIDGIQSLSFDIETNKKMVTAQELDYKVRKYLTDNKQLYTNGPSKYETGYIKFIPKNKE 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        151 VSIDGIQSLSFDIETNKKMVTAQELDYKVRKYLTDNKQLYTNGPSKYETGYIKFIPKNKE 210

Qy        181 SFWFDFFPEPEFTQSKYLMIYKDNETLDSNTSQIEVYLTTK 221
              |||||||||||||||||||||||||||||||||||||||||
Db        211 SFWFDFFPEPEFTQSKYLMIYKDNETLDSNTSQIEVYLTTK 251


RESULT 6
AAW12147
ID   AAW12147 standard; protein; 251 AA.
XX
AC   AAW12147;
XX
DT   04-NOV-1997  (first entry)
XX
DE   Streptococcus pyogenes Streptococcal toxin A mutant Asn20Asp.
XX
KW   Streptococcal; toxin A; SPE-A; non-lethal; mutant; production; vaccine;
KW   protection; treatment; cancer; neutralising antibody;
KW   streptococcal toxic shock syndrome; STSS; symptom; amelioration; fever;
KW   hypotension; group A streptococcal infection; myositis; fascitis;
KW   liver damage; T cell; lymphoma; ovarian; uterine.
XX
OS   Streptococcus pyogenes.
OS   Synthetic.
XX
FH   Key             Location/Qualifiers
FT   Peptide         1. .30
FT                   /label= sig_peptide
FT   Peptide         31. .251
FT                   /label= mat_peptide
FT   Misc-difference 50
FT                   /note= "wild type Asn replaced by Asp"
XX
PN   WO9640930-A1.
XX
PD   19-DEC-1996.
XX
PF   07-JUN-1996;   96WO-US010252.
XX
PR   07-JUN-1995;   95US-00480261.
XX
PA   (MINU ) UNIV MINNESOTA.
XX
PI   Schlievert PM,  Roggiani M,  Stoehr J,  Ohlendorf D;
XX
DR   WPI; 1997-099936/09.
XX
PT   Mutant SPE-A toxin with at least one amino acid change is substantially
PT   non-lethal - used in vaccine composition for treatment of cancer and
PT   streptococcal toxic shock syndrome etc.
XX
PS   Claim 5; Page; 102pp; English.
XX
CC   The present sequence is a non-lethal Streptococcus pyogenes Streptococcal
CC   toxin A (SPE-A) mutant, which can be used to produce vaccines to protect
CC   animals against wild type SPE-A and to treat cancer and streptococcal
CC   toxic shock syndrome (STSS). The mutant SPE-A causes neutralising
CC   antibodies (Ab) to be produced, which may be used to ameliorate STSS
CC   symptoms, e.g. fever, hypotension, group A streptococcal infection,
CC   myositis, fascitis and liver damage. The neutralising Ab is preferably
CC   administered in conjunction with antibiotic therapy. The mutant SPE-A is
CC   especially useful for treating T cell lymphomas, and ovarian and uterine
CC   cancer. It is thought that mutant SPE-A can be selectively toxic to T
CC   cell lymphoma cells. N.B. Sequence not given in the specification, but
CC   constructed using the wild type SPE-A sequence given on pages 77-79
XX
SQ   Sequence 251 AA;

  Query Match             99.6%;  Score 1173;  DB 2;  Length 251;
  Best Local Similarity   99.5%;  Pred. No. 2.4e-100;
  Matches  220;  Conservative    1;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 QQDPDPSQLHRSSLVKNLQNIYFLYEGDPVTHENVKSVDQLLSHHLIYNVSGPNYDKLKT 60
              ||||||||||||||||||:|||||||||||||||||||||||||||||||||||||||||||
Db         31 QQDPDPSQLHRSSLVKNLQDIYFLYEGDPVTHENVKSVDQLLSHHLIYNVSGPNYDKLKT 90

Qy         61 ELKNQEMATLFKDKNVDIYGVEYYHLCYLCENAERSACIYGGVTNHEGNHLEIPKKIVVK 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         91 ELKNQEMATLFKDKNVDIYGVEYYHLCYLCENAERSACIYGGVTNHEGNHLEIPKKIVVK 150

Qy        121 VSIDGIQSLSFDIETNKKMVTAQELDYKVRKYLTDNKQLYTNGPSKYETGYIKFIPKNKE 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        151 VSIDGIQSLSFDIETNKKMVTAQELDYKVRKYLTDNKQLYTNGPSKYETGYIKFIPKNKE 210

Qy        181 SFWFDFFPEPEFTQSKYLMIYKDNETLDSNTSQIEVYLTTK 221
              |||||||||||||||||||||||||||||||||||||||||
Db        211 SFWFDFFPEPEFTQSKYLMIYKDNETLDSNTSQIEVYLTTK 251
```

```
RESULT 7
AAB67344
ID   AAB67344 standard; peptide; 221 AA.
XX
AC   AAB67344;
XX
DT   23-APR-2001  (first entry)
XX
DE   Streptococcus pyrogenes toxin A protein.
XX
KW   Tumour; cancer; immune; enterotoxin.
XX
OS   Streptococcus pyogenes.
XX
PN   US6180097-B1.
XX
PD   30-JAN-2001.
XX
PF   30-OCT-1998;   98US-00183437.
XX
PR   03-OCT-1989;   89US-00416530.
PR   17-JAN-1990;   90US-00466577.
PR   17-JAN-1991;   91WO-US000342.
PR   01-JUN-1992;   92US-00891718.
PR   02-MAR-1993;   93US-00025144.
PR   31-JAN-1994;   94US-00189424.
PR   19-JUN-1995;   95US-00491746.
XX
PA   (TERM/) TERMAN D S.
XX
PI   Terman DS;
XX
DR   WPI; 2001-158657/16.
XX
PT   Tumor cell capable of stimulating antitumor immune reactivity in vitro or
PT   in vivo comprises exogenous nucleic acids encoding a superantigen and a
PT   costimulatory molecule.
XX
PS   Disclosure; Fig 2; 16pp; English.
XX
CC   The present invention relates to a tumour cell capable of stimulating
CC   antitumor immune reactivity in vitro or in vivo contains and expresses an
CC   exogenous nucleic acid molecule encoding a superantigen or its active
CC   fragment and an exogenous nucleic acid molecule encoding a costimulatory
CC   molecule that activates T cells in conjunction with an antigenic
CC   stimulus. The invention may be used for cancer therapy by stimulating an
CC   anticancer immune response in vivo or ex vivo
XX
SQ   Sequence 221 AA;

  Query Match             99.2%;  Score 1169;  DB 4;  Length 221;
  Best Local Similarity   99.5%;  Pred. No. 4.8e-100;
  Matches  220;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy          1 QQDPDPSQLHRSSLVKNLQNIYFLYEGDPVTHENVKSVDQLLSHHLIYNVSGPNYDKLKT 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 QQDPDPSQLHRSSLVKNLQNIYFLYEGDPVTHENVKSVDQLLSHDLIYNVSGPNYDKLKT 60

Qy         61 ELKNQEMATLFKDKNVDIYGVEYYHLCYLCENAERSACIYGGVTNHEGNHLEIPKKIVVK 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 ELKNQEMATLFKDKNVDIYGVEYYHLCYLCENAERSACIYGGVTNHEGNHLEIPKKIVVK 120

Qy        121 VSIDGIQSLSFDIETNKKMVTAQELDYKVRKYLTDNKQLYTNGPSKYETGYIKFIPKNKE 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 VSIDGIQSLSFDIETNKKMVTAQELDYKVRKYLTDNKQLYTNGPSKYETGYIKFIPKNKE 180

Qy        181 SFWFDFFPEPEFTQSKYLMIYKDNETLDSNTSQIEVYLTTK 221
              |||||||||||||||||||||||||||||||||||||||||
Db        181 SFWFDFFPEPEFTQSKYLMIYKDNETLDSNTSQIEVYLTTK 221


RESULT 8
AAW12148
ID   AAW12148 standard; protein; 251 AA.
XX
AC   AAW12148;
XX
DT   04-NOV-1997  (first entry)
XX
DE   Streptococcus pyogenes Streptococcal toxin A mutant N20D/K157E.
XX
KW   Streptococcal; toxin A; SPE-A; non-lethal; mutant; production; vaccine;
KW   protection; treatment; cancer; neutralising antibody;
KW   streptococcal toxic shock syndrome; STSS; symptom; amelioration; fever;
KW   hypotension; group A streptococcal infection; myositis; fascitis;
KW   liver damage; T cell; lymphoma; ovarian; uterine.
XX
OS   Streptococcus pyogenes.
OS   Synthetic.
XX
FH   Key             Location/Qualifiers
FT   Peptide         1. .30
FT                   /label= sig_peptide
FT   Peptide         31. .251
FT                   /label= mat_peptide
FT   Misc-difference 50
FT                   /note= "wild type Asn replaced by Asp"
FT   Misc-difference 187
FT                   /note= "wild type Lys replaced by Glu"
XX
PN   WO9640930-A1.
XX
PD   19-DEC-1996.
XX
PF   07-JUN-1996;   96WO-US010252.
XX
PR   07-JUN-1995;   95US-00480261.
XX
PA   (MINU ) UNIV MINNESOTA.
XX
PI   Schlievert PM,  Roggiani M,  Stoehr J,  Ohlendorf D;
XX
DR   WPI; 1997-099936/09.
XX
PT   Mutant SPE-A toxin with at least one amino acid change is substantially
PT   non-lethal - used in vaccine composition for treatment of cancer and
PT   streptococcal toxic shock syndrome etc.
XX
PS   Claim 5; Page; 102pp; English.
XX
CC   The present sequence is a non-lethal Streptococcus pyogenes Streptococcal
CC   toxin A (SPE-A) mutant, which can be used to produce vaccines to protect
CC   animals against wild type SPE-A and to treat cancer and streptococcal
CC   toxic shock syndrome (STSS). The mutant SPE-A causes neutralising
CC   antibodies (Ab) to be produced, which may be used to ameliorate STSS
CC   symptoms, e.g. fever, hypotension, group A streptococcal infection,
CC   myositis, fascitis and liver damage. The neutralising Ab is preferably
CC   administered in conjunction with antibiotic therapy. The mutant SPE-A is
CC   especially useful for treating T cell lymphomas, and ovarian and uterine
CC   cancer. It is thought that mutant SPE-A can be selectively toxic to T
CC   cell lymphoma cells. N.B. Sequence not given in the specification, but
CC   constructed using the wild type SPE-A sequence given on pages 77-79
XX
SQ   Sequence 251 AA;

  Query Match             99.2%;  Score 1169;  DB 2;  Length 251;
  Best Local Similarity   99.1%;  Pred. No. 5.7e-100;
  Matches  219;  Conservative    2;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 QQDPDPSQLHRSSLVKNLQNIYFLYEGDPVTHENVKSVDQLLSHHLIYNVSGPNYDKLKT 60
              |||||||||||||||||||:||||||||||||||||||||||||||||||||||||||||
Db         31 QQDPDPSQLHRSSLVKNLQDIYFLYEGDPVTHENVKSVDQLLSHHLIYNVSGPNYDKLKT 90

Qy         61 ELKNQEMATLFKDKNVDIYGVEYYHLCYLCENAERSACIYGGVTNHEGNHLEIPKKIVVK 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         91 ELKNQEMATLFKDKNVDIYGVEYYHLCYLCENAERSACIYGGVTNHEGNHLEIPKKIVVK 150
```

```
Qy        121 VSIDGIQSLSFDIETNKKMVTAQELDYKVRKYLTDNKQLYTNGPSKYETGYIKFIPKNKE 180
              ||||||||||||||||||||||||||||||||||||:||||||||||||||||||||||
Db        151 VSIDGIQSLSFDIETNKKMVTAQELDYKVRKYLTDNEQLYTNGPSKYETGYIKFIPKNKE 210

Qy        181 SFWFDFFPEPEFTQSKYLMIYKDNETLDSNTSQIEVYLTTK 221
              |||||||||||||||||||||||||||||||||||||||||
Db        211 SFWFDFFPEPEFTQSKYLMIYKDNETLDSNTSQIEVYLTTK 251
```

RESULT 9
AAW12153
ID AAW12153 standard; protein; 251 AA.
XX
AC AAW12153;
XX
DT 04-NOV-1997 (first entry)
XX
DE Streptococcus pyogenes Streptococcal toxin A mutant Cys90Ser.
XX
KW Streptococcal; toxin A; SPE-A; non-lethal; mutant; production; vaccine;
KW protection; treatment; cancer; neutralising antibody;
KW streptococcal toxic shock syndrome; STSS; symptom; amelioration; fever;
KW hypotension; group A streptococcal infection; myositis; fascitis;
KW liver damage; T cell; lymphoma; ovarian; uterine.
XX
OS Streptococcus pyogenes.
OS Synthetic.
XX
FH Key Location/Qualifiers
FT Peptide 1. .30
FT /label= sig_peptide
FT Peptide 31. .251
FT /label= mat_peptide
FT Misc-difference 120
FT /note= "wild type Cys replaced by Ser"
XX
PN WO9640930-A1.
XX
PD 19-DEC-1996.
XX
PF 07-JUN-1996; 96WO-US010252.
XX
PR 07-JUN-1995; 95US-00480261.
XX
PA (MINU) UNIV MINNESOTA.
XX
PI Schlievert PM, Roggiani M, Stoehr J, Ohlendorf D;
XX
DR WPI; 1997-099936/09.
XX
PT Mutant SPE-A toxin with at least one amino acid change is substantially
PT non-lethal - used in vaccine composition for treatment of cancer and
PT streptococcal toxic shock syndrome etc.
XX
PS Example 4; Page; 102pp; English.
XX
CC The present sequence is a non-lethal Streptococcus pyogenes Streptococcal
CC toxin A (SPE-A) mutant, which can be used to produce vaccines to protect
CC animals against wild type SPE-A and to treat cancer and streptococcal
CC toxic shock syndrome (STSS). The mutant SPE-A causes neutralising
CC antibodies (Ab) to be produced, which may be used to ameliorate STSS
CC symptoms, e.g. fever, hypotension, group A streptococcal infection,
CC myositis, fascitis and liver damage. The neutralising Ab is preferably
CC administered in conjunction with antibiotic therapy. The mutant SPE-A is
CC especially useful for treating T cell lymphomas, and ovarian and uterine
CC cancer. It is thought that mutant SPE-A can be selectively toxic to T
CC cell lymphoma cells. N.B. Sequence not given in the specification, but
CC constructed using the wild type SPE-A sequence given on pages 77-79
XX
SQ Sequence 251 AA;

Query Match 99.2%; Score 1168; DB 2; Length 251;
Best Local Similarity 99.5%; Pred. No. 7e-100;
Matches 220; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

```
Qy          1 QQDPDPSQLHRSSLVKNLQNIYFLYEGDPVTHENVKSVDQLLSHHLIYNVSGPNYDKLKT 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         31 QQDPDPSQLHRSSLVKNLQNIYFLYEGDPVTHENVKSVDQLLSHHLIYNVSGPNYDKLKT 90

Qy         61 ELKNQEMATLFKDKNVDIYGVEYYHLCYLCENAERSACIYGGVTNHEGNHLEIPKKIVVK 120
              |||||||||||||||||||||||||| ||||||||||||||||||||||||||||||||||
Db         91 ELKNQEMATLFKDKNVDIYGVEYYHLCYLSENAERSACIYGGVTNHEGNHLEIPKKIVVK 150

Qy        121 VSIDGIQSLSFDIETNKKMVTAQELDYKVRKYLTDNKQLYTNGPSKYETGYIKFIPKNKE 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        151 VSIDGIQSLSFDIETNKKMVTAQELDYKVRKYLTDNKQLYTNGPSKYETGYIKFIPKNKE 210

Qy        181 SFWFDFFPEPEFTQSKYLMIYKDNETLDSNTSQIEVYLTTK 221
              |||||||||||||||||||||||||||||||||||||||||
Db        211 SFWFDFFPEPEFTQSKYLMIYKDNETLDSNTSQIEVYLTTK 251
```

RESULT 10
AAW12151
ID AAW12151 standard; protein; 251 AA.
XX
AC AAW12151;
XX
DT 04-NOV-1997 (first entry)
XX
DE Streptococcus pyogenes Streptococcal toxin A mutant Cys87Ser.
XX
KW Streptococcal; toxin A; SPE-A; non-lethal; mutant; production; vaccine;
KW protection; treatment; cancer; neutralising antibody;
KW streptococcal toxic shock syndrome; STSS; symptom; amelioration; fever;
KW hypotension; group A streptococcal infection; myositis; fascitis;
KW liver damage; T cell; lymphoma; ovarian; uterine.
XX
OS Streptococcus pyogenes.
OS Synthetic.
XX
FH Key Location/Qualifiers
FT Peptide 1. .30
FT /label= sig_peptide
FT Peptide 31. .251
FT /label= mat_peptide
FT Misc-difference 117
FT /note= "wild type Cys replaced by Ser"
XX
PN WO9640930-A1.
XX
PD 19-DEC-1996.
XX
PF 07-JUN-1996; 96WO-US010252.
XX
PR 07-JUN-1995; 95US-00480261.
XX
PA (MINU) UNIV MINNESOTA.
XX
PI Schlievert PM, Roggiani M, Stoehr J, Ohlendorf D;
XX
DR WPI; 1997-099936/09.
XX
PT Mutant SPE-A toxin with at least one amino acid change is substantially
PT non-lethal - used in vaccine composition for treatment of cancer and
PT streptococcal toxic shock syndrome etc.
XX
PS Example 4; Page; 102pp; English.
XX
CC The present sequence is a non-lethal Streptococcus pyogenes Streptococcal
CC toxin A (SPE-A) mutant, which can be used to produce vaccines to protect
CC animals against wild type SPE-A and to treat cancer and streptococcal
CC toxic shock syndrome (STSS). The mutant SPE-A causes neutralising
CC antibodies (Ab) to be produced, which may be used to ameliorate STSS
CC symptoms, e.g. fever, hypotension, group A streptococcal infection,

```
CC   myositis, fascitis and liver damage. The neutralising Ab is preferably
CC   administered in conjunction with antibiotic therapy. The mutant SPE-A is
CC   especially useful for treating T cell lymphomas, and ovarian and uterine
CC   cancer. It is thought that mutant SPE-A can be selectively toxic to T
CC   cell lymphoma cells. N.B. Sequence not given in the specification, but
CC   constructed using the wild type SPE-A sequence given on pages 77-79
XX
SQ   Sequence 251 AA;

  Query Match             99.2%;  Score 1168;  DB 2;  Length 251;
  Best Local Similarity   99.5%;  Pred. No. 7e-100;
  Matches  220;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy          1 QQDPDPSQLHRSSLVKNLQNIYFLYEGDPVTHENVKSVDQLLSHHLIYNVSGPNYDKLKT 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         31 QQDPDPSQLHRSSLVKNLQNIYFLYEGDPVTHENVKSVDQLLSHHLIYNVSGPNYDKLKT 90

Qy         61 ELKNQEMATLFKDKNVDIYGVEYYHLCYLCENAERSACIYGGVTNHEGNHLEIPKKIVVK 120
              |||||||||||||||||||||||||| ||||||||||||||||||||||||||||||||
Db         91 ELKNQEMATLFKDKNVDIYGVEYYHLSYLCENAERSACIYGGVTNHEGNHLEIPKKIVVK 150

Qy        121 VSIDGIQSLSFDIETNKKMVTAQELDYKVRKYLTDNKQLYTNGPSKYETGYIKFIPKNKE 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        151 VSIDGIQSLSFDIETNKKMVTAQELDYKVRKYLTDNKQLYTNGPSKYETGYIKFIPKNKE 210

Qy        181 SFWFDFFPEPEFTQSKYLMIYKDNETLDSNTSQIEVYLTTK 221
              |||||||||||||||||||||||||||||||||||||||||
Db        211 SFWFDFFPEPEFTQSKYLMIYKDNETLDSNTSQIEVYLTTK 251


RESULT 11
AAW12152
ID   AAW12152 standard; protein; 251 AA.
XX
AC   AAW12152;
XX
DT   04-NOV-1997  (first entry)
XX
DE   Streptococcus pyogenes Streptococcal toxin A mutant Cys98Ser.
XX
KW   Streptococcal; toxin A; SPE-A; non-lethal; mutant; production; vaccine;
KW   protection; treatment; cancer; neutralising antibody;
KW   streptococcal toxic shock syndrome; STSS; symptom; amelioration; fever;
KW   hypotension; group A streptococcal infection; myositis; fascitis;
KW   liver damage; T cell; lymphoma; ovarian; uterine.
XX
OS   Streptococcus pyogenes.
OS   Synthetic.
XX
FH   Key             Location/Qualifiers
FT   Peptide         1. .30
FT                   /label= sig_peptide
FT   Peptide         31. .251
FT                   /label= mat_peptide
FT   Misc-difference 128
FT                   /note= "wild type Cys replaced by Ser"
XX
PN   WO9640930-A1.
XX
PD   19-DEC-1996.
XX
PF   07-JUN-1996;   96WO-US010252.
XX
PR   07-JUN-1995;   95US-00480261.
XX
PA   (MINU ) UNIV MINNESOTA.
XX
PI   Schlievert PM,  Roggiani M,  Stoehr J,  Ohlendorf D;
XX
DR   WPI; 1997-099936/09.
XX
PT   Mutant SPE-A toxin with at least one amino acid change is substantially
PT   non-lethal - used in vaccine composition for treatment of cancer and
PT   streptococcal toxic shock syndrome etc.
XX
PS   Example 4; Page; 102pp; English.
XX
CC   The present sequence is a non-lethal Streptococcus pyogenes Streptococcal
CC   toxin A (SPE-A) mutant, which can be used to produce vaccines to protect
CC   animals against wild type SPE-A and to treat cancer and streptococcal
CC   toxic shock syndrome (STSS). The mutant SPE-A causes neutralising
CC   antibodies (Ab) to be produced, which may be used to ameliorate STSS
CC   symptoms, e.g. fever, hypotension, group A streptococcal infection,
CC   myositis, fascitis and liver damage. The neutralising Ab is preferably
CC   administered in conjunction with antibiotic therapy. The mutant SPE-A is
CC   especially useful for treating T cell lymphomas, and ovarian and uterine
CC   cancer. It is thought that mutant SPE-A can be selectively toxic to T
CC   cell lymphoma cells. N.B. Sequence not given in the specification, but
CC   constructed using the wild type SPE-A sequence given on pages 77-79
XX
SQ   Sequence 251 AA;

  Query Match             99.2%;  Score 1168;  DB 2;  Length 251;
  Best Local Similarity   99.5%;  Pred. No. 7e-100;
  Matches  220;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy          1 QQDPDPSQLHRSSLVKNLQNIYFLYEGDPVTHENVKSVDQLLSHHLIYNVSGPNYDKLKT 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         31 QQDPDPSQLHRSSLVKNLQNIYFLYEGDPVTHENVKSVDQLLSHHLIYNVSGPNYDKLKT 90

Qy         61 ELKNQEMATLFKDKNVDIYGVEYYHLCYLCENAERSACIYGGVTNHEGNHLEIPKKIVVK 120
              |||||||||||||||||||||||||||||||| ||||||||||||||||||||||||||||
Db         91 ELKNQEMATLFKDKNVDIYGVEYYHLCYLCENAERSASIYGGVTNHEGNHLEIPKKIVVK 150

Qy        121 VSIDGIQSLSFDIETNKKMVTAQELDYKVRKYLTDNKQLYTNGPSKYETGYIKFIPKNKE 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        151 VSIDGIQSLSFDIETNKKMVTAQELDYKVRKYLTDNKQLYTNGPSKYETGYIKFIPKNKE 210

Qy        181 SFWFDFFPEPEFTQSKYLMIYKDNETLDSNTSQIEVYLTTK 221
              |||||||||||||||||||||||||||||||||||||||||
Db        211 SFWFDFFPEPEFTQSKYLMIYKDNETLDSNTSQIEVYLTTK 251


RESULT 12
ABU62460
ID   ABU62460 standard; protein; 251 AA.
XX
AC   ABU62460;
XX
DT   27-AUG-2003  (first entry)
XX
DE   Streptococcus pyrogenic toxin a L42A mutant.
XX
KW   SPEa; streptococcus pyrogenic enterotoxin a; mutant; vaccine; mutein;
KW   superantigen toxin; MHC; superantigen-associated bacterial infection;
KW   bacterial infection; antibacterial.
XX
OS   Streptococcus sp.
OS   Synthetic.
XX
FH   Key             Location/Qualifiers
FT   Peptide         1. .30
FT                   /label= Signal_peptide
FT   Protein         31. .251
FT                   /label= Mature_SPEa_L42R
FT   Misc-difference 72
FT                   /note= "Wild-type Leu substituted by Ala"
XX
PN   US2003036644-A1.
XX
PD   20-FEB-2003.
XX
PF   26-NOV-2001; 2001US-00002784.
XX
PR   25-JUN-1997;   97US-00882431.
PR   01-SEP-1998;   98US-00144776.
```

XX
PA (ULRI/) ULRICH R G.
XX
PI Ulrich RG;
XX
DR WPI; 2003-492125/46.
XX
PT New superantigen toxin DNA fragment, useful for preparing a composition
PT for treating or preventing bacterial infection.
XX
PS Example 13; Page; 68pp; English.
XX
CC The invention relates to an isolated and purified superantigen toxin DNA
CC fragment is altered so that binding of the encoded altered toxin to
CC either the MHC class II or T cell antigen receptor is altered. Also
CC included are a recombinant DNA construct (comprising a vector and an
CC isolated and purified altered superantigen toxin DNA fragment), a host
CC cell transformed with the recombinant DNA construct, producing altered
CC superantigen toxin, an altered TSST-1 (toxic shock syndrome toxin)
CC superantigen toxin peptide, diagnosing superantigen-associated bacterial
CC infection, a vaccine (comprising an altered superantigen toxin for
CC producing antigenic and immunogenic response resulting in the protection
CC of a mammal against superantigen-associated bacterial infection),
CC treating/ameliorating a superantigen-associated bacterial infection, an
CC antisera isolated from individuals immunised with one or more altered
CC TSST-1 superantigen toxin and an antibody which recognises altered TSST-
CC 1. Also disclosed are mutated staphylococcal enterotoxin A, B, C1 (SEA,
CC SEB, SEC1) and streptococcal pyrogenic enterotoxin a and b (SPEa and
CC SPEb). The superantigen toxin DNA fragment is useful for preparing a
CC composition for treating or preventing bacterial infection. The present
CC sequence represents the L42A (with reference to the mature protein)
CC mutant of SPEa. Note: The present sequence is not shown in the
CC specification but was created by the indexer using the wild-type sequence
CC and the information in the specification
XX
SQ Sequence 251 AA;

Not prior

Query Match 98.8%; Score 1164; DB 7; Length 251;
Best Local Similarity 99.1%; Pred. No. 1.6e-99;
Matches 219; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

```
Qy          1 QQDPDPSQLHRSSLVKNLQNIYFLYEGDPVTHENVKSVDQLLSHHLIYNVSGPNYDKLKT 60
              |||||||||||||||||||||||||||||||||||||||||| || ||||||||||||||||
Db         31 QQDPDPSQLHRSSLVKNLQNIYFLYEGDPVTHENVKSVDQLASHDLIYNVSGPNYDKLKT 90

Qy         61 ELKNQEMATLFKDKNVDIYGVEYYHLCYLCENAERSACIYGGVTNHEGNHLEIPKKIVVK 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         91 ELKNQEMATLFKDKNVDIYGVEYYHLCYLCENAERSACIYGGVTNHEGNHLEIPKKIVVK 150

Qy        121 VSIDGIQSLSFDIETNKKMVTAQELDYKVRKYLTDNKQLYTNGPSKYETGYIKFIPKNKE 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        151 VSIDGIQSLSFDIETNKKMVTAQELDYKVRKYLTDNKQLYTNGPSKYETGYIKFIPKNKE 210

Qy        181 SFWFDFFPEPEFTQSKYLMIYKDNETLDSNTSQIEVYLTTK 221
              |||||||||||||||||||||||||||||||||||||||||
Db        211 SFWFDFFPEPEFTQSKYLMIYKDNETLDSNTSQIEVYLTTK 251
```

RESULT 13
AAR13209
ID AAR13209 standard; protein; 221 AA.
XX
AC AAR13209;
XX
DT 27-AUG-2003 (revised)
DT 15-OCT-1991 (first entry)
XX
DE Streptococcal pyrogenic enterotoxin A.
XX
KW SPE A; cancer treatment; pyrogen; tumouricide; scarlet fever.
XX
OS Streptococcus sp.
XX
PN WO9110680-A.
XX
PD 25-JUL-1991.
XX
PF 17-JAN-1990; 90US-00466577.
XX
PR 17-JAN-1990; 90US-00466577.
XX
PA (TERM/) TERMAN D S.
XX
PI Terman DS;
XX
DR WPI; 1991-237984/32.
XX
PT Treating cancer with enterotoxin from Staphylococcus aureus -
PT administered by IV injection, having same tumoricidal activity as
PT Staphylococcal protein A without potential toxic reactions.
XX
PS Disclosure; Fig 1; 74pp; English.
XX
CC SPE A can be used for tumouricidal treatment, esp. with a haemolysin.
CC Synthetic polypeptides having structural homology to Streptococcal
CC pyrogenic exotoxins are claimed, provided the homology includes
CC statistically significant sequence homology, alignment of Cysteine
CC residues and similar hydropathy profiles. See AAR13203-R13211. (Updated
CC on 27-AUG-2003 to correct OS field.)
XX
SQ Sequence 221 AA;

Query Match 98.7%; Score 1163; DB 2; Length 221;
Best Local Similarity 99.1%; Pred. No. 1.7e-99;
Matches 219; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

```
Qy          1 QQDPDPSQLHRSSLVKNLQNIYFLYEGDPVTHENVKSVDQLLSHHLIYNVSGPNYDKLKT 60
              |||||||||||||||||||||||||||||||||||||||||| ||||||||||||||||
Db          1 QQDPDPSQLHRSSLVKNLQNIYFLYEGDPVTHENVKSVDQLLSHDLIYNVSGPNYDKLKT 60

Qy         61 ELKNQEMATLFKDKNVDIYGVEYYHLCYLCENAERSACIYGGVTNHEGNHLEIPKKIVVK 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 ELKNQEMATLFKDKNVDIYGVEYYHLCYLCENAERSACIYGGVTNHEGNHLEIPKKIVVK 120

Qy        121 VSIDGIQSLSFDIETNKKMVTAQELDYKVRKYLTDNKQLYTNGPSKYETGYIKFIPKNKE 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 VSIDGIQSLSFDIETNKKMVTAQELDYKVRKYLTDNKQLYTNGPSKYETGYIKFIPKNKE 180

Qy        181 SFWFDFFPEPEFTQSKYLMIYKDNETLDSNTSQIEVYLTTK 221
              ||||| |||||||||||||||||||||||||||||||||||
Db        181 SFWFDLFPEPEFTQSKYLMIYKDNETLDSNTSQIEVYLTTK 221
```

asp acid
@ posn 45

RESULT 14
AAR45017
ID AAR45017 standard; protein; 221 AA.
XX
AC AAR45017;
XX
DT 25-MAR-2003 (revised)
DT 08-JUN-1994 (first entry)
XX
DE Staphylococcal enterotoxin SPE A.
XX
KW Staphylococcal enterotoxin; SE; cancer; tumouricidal agent;
KW autoimmune disease; toxicity; Protein A; perfusion system.
XX
OS Staphylococcus aureus.
XX
PN WO9324136-A1.
XX
PD 09-DEC-1993.
XX
PF 01-JUN-1993; 93WO-US005213.
XX
PR 01-JUN-1992; 92US-00891718.

Not asparagine (N)!

XX
PA (TERM/) TERMAN D S.
PA (STON/) STONE J L.
XX
PI Terman DS, Stone JL;
XX
DR WPI; 1993-405418/50.
XX
PT Use of staphylococcal enterotoxin(s) and homologues - for treating cancer
PT in a patient or for the treatment of auto-immune diseases.
XX
PS Disclosure; Fig 1; 90pp; English.
XX
CC The sequences given in AAR45011-21 are Staphylococcal enterotoxins (SEs)
CC which may be used in the methods of the invention for treating cancer in
CC a patient. These SEs, and homologues of them, can be used as tumouricidal
CC agents for treating cancers and autoimmune disease. They exhibit
CC tumouricidal activity and toxicity identical to that observed for the
CC Protein A perfusion system. They may be administered by i.v. injection.
CC (Updated on 25-MAR-2003 to correct PN field.)
XX
SQ Sequence 221 AA;

Query Match 98.7%; Score 1163; DB 2; Length 221;
Best Local Similarity 99.1%; Pred. No. 1.7e-99;
Matches 219; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

```
Qy          1 QQDPDPSQLHRSSLVKNLQNIYFLYEGDPVTHENVKSVDQLLSHHLIYNVSGPNYDKLKT 60
              ||||||||||||||||||||||||||||||||||||||||||| ||||||||||||||||
Db          1 QQDPDPSQLHRSSLVKNLQNIYFLYEGDPVTHENVKSVDQLLSHDLIYNVSGPNYDKLKT 60

Qy         61 ELKNQEMATLFKDKNVDIYGVEYYHLCYLCENAERSACIYGGVTNHEGNHLEIPKKIVVK 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 ELKNQEMATLFKDKNVDIYGVEYYHLCYLCENAERSACIYGGVTNHEGNHLEIPKKIVVK 120

Qy        121 VSIDGIQSLSFDIETNKKMVTAQELDYKVRKYLTDNKQLYTNGPSKYETGYIKFIPKNKE 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 VSIDGIQSLSFDIETNKKMVTAQELDYKVRKYLTDNKQLYTNGPSKYETGYIKFIPKNKE 180

Qy        181 SFWFDFFPEPEFTQSKYLMIYKDNETLDSNTSQIEVYLTTK 221
              ||||| |||||||||||||||||||||||||||||||||||
Db        181 SFWFDLFPEPEFTQSKYLMIYKDNETLDSNTSQIEVYLTTK 221
```

RESULT 15
ABB76240
ID ABB76240 standard; protein; 221 AA.
XX
AC ABB76240;
XX
DT 09-AUG-2002 (first entry)
XX
DE Staphylococcus pyogenes exotoxin A.
XX
KW Exotoxin A; SPE A; superantigen; antigen; tumour; cancer; antitumour;
KW therapy.
XX
OS Streptococcus pyogenes.
XX
PN US2002051765-A1.
XX
PD 02-MAY-2002.
XX
PF 19-DEC-2000; 2000US-00741503.
XX
PR 03-OCT-1989; 89US-00416530.
PR 17-JAN-1990; 90US-00466577.
PR 17-JAN-1991; 91WO-US000342.
PR 01-JUN-1992; 92US-00891718.
PR 02-MAR-1993; 93US-00025144.
PR 31-JAN-1994; 94US-00189424.
PR 19-JUN-1995; 95US-00491746.
XX

PA (TERM/) TERMAN D S.
XX
PI Terman DS;
XX
DR WPI; 2002-415198/44.
XX
PT Reagent for treating cancer without the need for e.g. radiotherapy,
PT comprises a specific V beta subset of T cells sensitized to a growing
PT tumor and stimulated with superantigens.
XX
PS Disclosure; Fig 2; 17pp; English.
XX
CC The present sequence is the protein sequence of exotoxin A (SPE A) of
CC Streptococcus pyogenes. Similarity is shown, in several stretches of
CC sequence, between staphylococcal enterotoxins, streptococcal pyrogenic
CC exotoxins and staphylococcal exfoliative toxins (see ABB76234-44). In the
CC present invention, synthetic polypeptides useful in tumour therapy and in
CC blocking or destroying autoreactive T and B lymphocyte populations are
CC characterised by substantial structural homology to staphylococcal
CC enterotoxin A and enterotoxin B, and to streptococcal pyrogenic
CC exotoxins, with statistically significant sequence homology and
CC similarity (Z value of Lipman and Pearson algorithm in Monte Carlo
CC analysis exceeding 6) to include alignment of cysteine residues and
CC similar hydropathy profiles. These superantigens are used to treat solid
CC tumours, including their metastases, without radiation, surgery or
CC standard chemotherapeutic agents. A claimed method of human cancer
CC treatment involves contacting haematopoietic cells from a patient with
CC one or more superantigens ex vivo to generate stimulated cells, selecting
CC a specific V beta subset of cells, and reintroducing these cells into the
CC patient to induce an in vivo therapeutic, tumouricidal reaction
XX
SQ Sequence 221 AA;

Query Match 98.7%; Score 1163; DB 5; Length 221;
Best Local Similarity 99.1%; Pred. No. 1.7e-99;
Matches 219; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

```
Qy          1 QQDPDPSQLHRSSLVKNLQNIYFLYEGDPVTHENVKSVDQLLSHHLIYNVSGPNYDKLKT 60
              ||||||||||||||||||||||||||||||||||||||||||| ||||||||||||||||
Db          1 QQDPDPSQLHRSSLVKNLQNIYFLYEGDPVTHENVKSVDQLLSHDLIYNVSGPNYDKLKT 60

Qy         61 ELKNQEMATLFKDKNVDIYGVEYYHLCYLCENAERSACIYGGVTNHEGNHLEIPKKIVVK 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 ELKNQEMATLFKDKNVDIYGVEYYHLCYLCENAERSACIYGGVTNHEGNHLEIPKKIVVK 120

Qy        121 VSIDGIQSLSFDIETNKKMVTAQELDYKVRKYLTDNKQLYTNGPSKYETGYIKFIPKNKE 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 VSIDGIQSLSFDIETNKKMVTAQELDYKVRKYLTDNKQLYTNGPSKYETGYIKFIPKNKE 180

Qy        181 SFWFDFFPEPEFTQSKYLMIYKDNETLDSNTSQIEVYLTTK 221
              ||||| |||||||||||||||||||||||||||||||||||
Db        181 SFWFDLFPEPEFTQSKYLMIYKDNETLDSNTSQIEVYLTTK 221
```

Search completed: May 7, 2004, 12:07:38
Job time : 62 secs

no substitutes

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: May 9, 2004, 06:36:07 ; Search time 4334 Seconds
(without alignments)
12753.783 Million cell updates/sec

Title: US-10-625-221-12
Perfect score: 1851
Sequence: 1 ccatcacgcatcactcatgt.....ttagcaactattttatcgtc 1851

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 27513289 seqs, 14931090276 residues

Total number of hits satisfying chosen parameters: 55026578

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : EST:*
1: em_estba:*
2: em_esthum:*
3: em_estin:*
4: em_estmu:*
5: em_estov:*
6: em_estpl:*
7: em_estro:*
8: em_htc:*
9: gb_est1:*
10: gb_est2:*
11: gb_htc:*
12: gb_est3:*
13: gb_est4:*
14: gb_est5:*
15: em_estfun:*
16: em_estom:*
17: em_gss_hum:*
18: em_gss_inv:*
19: em_gss_pln:*
20: em_gss_vrt:*
21: em_gss_fun:*
22: em_gss_mam:*
23: em_gss_mus:*
24: em_gss_pro:*
25: em_gss_rod:*
26: em_gss_phg:*
27: em_gss_vrl:*
28: gb_gss1:*
29: gb_gss2:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.		Score	% Query Match	Length	DB	ID	Description
c	1	97	5.2	1200	13	BX415878	BX415878 BX415878
c	2	96.8	5.2	1200	13	BX437758	BX437758 BX437758
	3	96.2	5.2	1277	28	CC253231	CC253231 CH261-180
	4	93.8	5.1	1101	29	CNS0039G	AL063921 Drosophil
c	5	91.6	4.9	1200	13	BX437739	BX437739 BX437739
c	6	91.4	4.9	1201	13	BX355654	BX355654 BX355654
c	7	91.2	4.9	1201	13	BX443774	BX443774 BX443774
c	8	90.2	4.9	1202	28	CC262481	CC262481 CH261-167
c	9	89.8	4.9	1124	13	BX436282	BX436282 BX436282
	10	89.8	4.9	1201	9	AL565455	AL565455 AL565455
c	11	89.6	4.8	1201	13	BX446296	BX446296 BX446296
c	12	88.8	4.8	1201	9	AL565455	AL565455 AL565455
c	13	88.4	4.8	1101	29	CNS0039G	AL063921 Drosophil
c	14	87.8	4.7	1101	29	CNS00EVL	AL069706 Drosophil
	15	87.8	4.7	1201	9	AL536104	AL536104 AL536104
c	16	87.8	4.7	1201	13	BX439779	BX439779 BX439779
	17	86.8	4.7	1201	13	BX458623	BX458623 BX458623
c	18	86.6	4.7	1896	29	CG753083	CG753083 P048-1-C0
c	19	86.4	4.7	1056	13	BX415058	BX415058 BX415058
c	20	86.4	4.7	1592	29	CG750135	CG750135 P044-3-D0
c	21	86.2	4.7	945	29	CNS04D0K	AL285149 Tetraodon
	22	86.2	4.7	1200	13	BX437758	BX437758 BX437758
c	23	86.2	4.7	1201	9	AL536104	AL536104 AL536104
c	24	85.8	4.6	1101	29	CNS0006J	AL062049 Drosophil
	25	85.2	4.6	1101	29	CNS00EVL	AL069706 Drosophil
	26	85.2	4.6	1200	13	BX415878	BX415878 BX415878
	27	85	4.6	1099	13	BX456575	BX456575 BX456575
	28	84.8	4.6	1187	28	B11102	B11102 F19C22-T7 I
c	29	84.6	4.6	1201	13	BX420717	BX420717 BX420717
c	30	83.2	4.5	1101	29	CNS003BD	AL064091 Drosophil
	31	82	4.4	961	29	CNS008HI	AL051882 Drosophil
c	32	81.6	4.4	639	29	CNS017QD	AL108367 Drosophil
c	33	81.4	4.4	899	13	BX453223	BX453223 BX453223
c	34	81.2	4.4	1201	13	BX355642	BX355642 BX355642
	35	80.8	4.4	1056	13	BX415058	BX415058 BX415058
c	36	80	4.3	1076	13	BX338020	BX338020 BX338020
c	37	80	4.3	1391	29	CG754863	CG754863 P050-2-G0
	38	79.4	4.3	1001	29	CNS0064G	AL062781 Drosophil
	39	79.4	4.3	1204	29	CNS016E2	AL106628 Drosophil
c	40	79.2	4.3	1101	29	CNS003B4	AL064082 Drosophil
c	41	79.2	4.3	1101	29	CNS016LI	AL106896 Drosophil
	42	79	4.3	1044	13	BX415231	BX415231 BX415231
c	43	78.8	4.3	646	14	CF547155	CF547155 AGENCOURT
	44	78.8	4.3	987	29	CNS014PQ	AL104456 Drosophil
	45	78.8	4.3	1200	13	BX437739	BX437739 BX437739

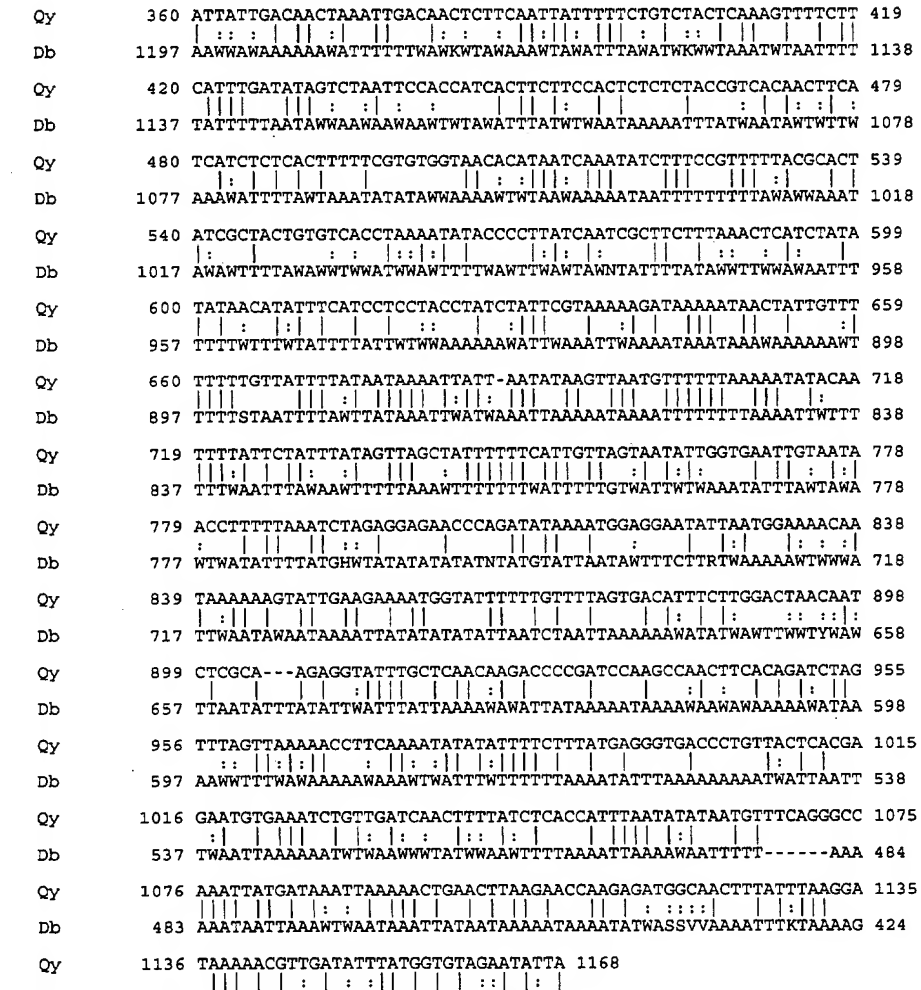
ALIGNMENTS

RESULT 1
BX415878/c
LOCUS BX415878 1200 bp mRNA linear EST 15-MAY-2003
DEFINITION BX415878 Homo sapiens THYMUS Homo sapiens cDNA clone CS0CAP008YI04 5-PRIME, mRNA sequence.
ACCESSION BX415878
VERSION BX415878.1 GI:30765550
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 1200)
AUTHORS Li,W.B., Gruber,C., Jessee,J. and Polayes,D.
TITLE Full-length cDNA libraries and normalization
JOURNAL Unpublished (2001)
COMMENT Contact: Genoscope
Genoscope - Centre National de Sequencage
BP 191 91006 EVRY cedex - France
Email: seqref@genoscope.cns.fr, Web : www.genoscope.cns.fr
Library was constructed by Life Technologies, a division of Invitrogen. Contact : Feng Liang Email : fliang@lifetech.com URL : http://fulllength.invitrogen.com/ InVitroGen Corporation 1600 Faraday Avenue Genoscope sequence ID : CS0CAP008BE02QP1.
FEATURES Location/Qualifiers
source 1. .1200
/organism="Homo sapiens"

/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="CS0CAP008YI04"
/tissue_type="THYMUS"
/clone_lib="Homo sapiens THYMUS"
/note="Vector: pCMVSPORT_6; 1st strand cDNA was primed with a NotI-oligo(dT) primer. Five prime end enriched, double-strand cDNA was digested with Not I and cloned into the Not I and EcoRV sites of the pCMVSPORT 6 vector. Library was not normalized."

ORIGIN

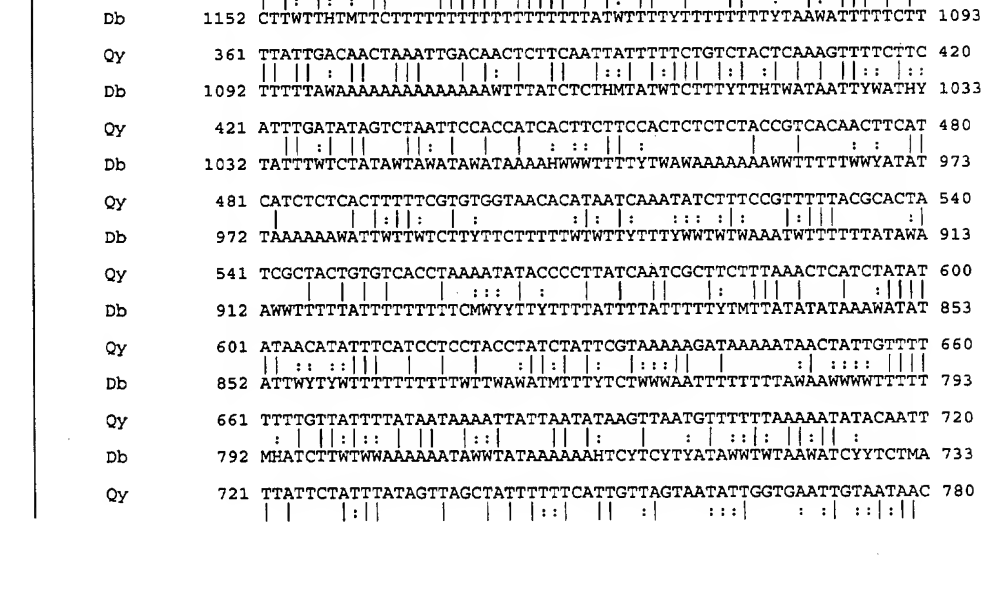
Query Match 5.2%; Score 97; DB 13; Length 1200;
Best Local Similarity 37.4%; Pred. No. 7.7e-06;
Matches 304; Conservative 111; Mismatches 388; Indels 10; Gaps 3;

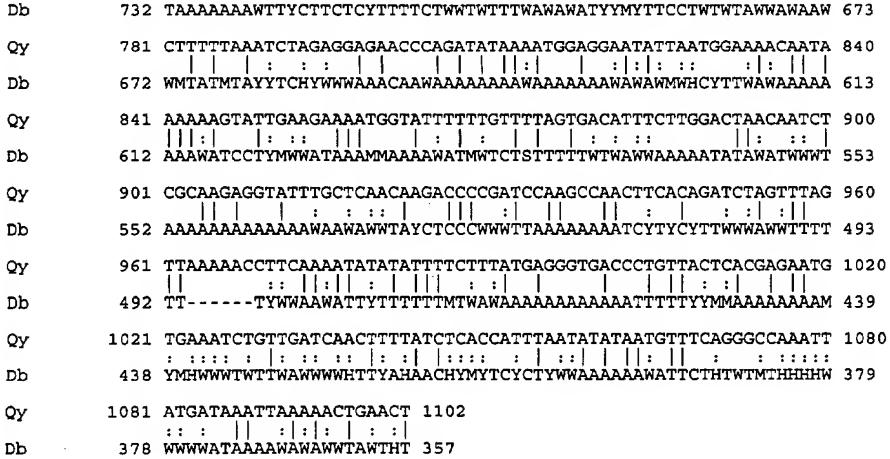


RESULT 2
BX437758/c
LOCUS BX437758 1200 bp mRNA linear EST 15-MAY-2003
DEFINITION BX437758 Homo sapiens THYMUS Homo sapiens cDNA clone CS0CAP008YB01 5-PRIME, mRNA sequence.
ACCESSION BX437758
VERSION BX437758.1 GI:30773605
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 1200)
AUTHORS Li,W.B., Gruber,C., Jessee,J. and Polayes,D.
TITLE Full-length cDNA libraries and normalization
JOURNAL Unpublished (2001)
COMMENT Contact: Genoscope
Genoscope - Centre National de Sequencage
BP 191 91006 EVRY cedex - France
Email: seqref@genoscope.cns.fr, Web : www.genoscope.cns.fr
Library was constructed by Life Technologies, a division of Invitrogen. Contact : Feng Liang Email : fliang@lifetech.com URL : http://fulllength.invitrogen.com/ InVitroGen Corporation 1600 Faraday Avenue Genoscope sequence ID : CS0CAP008CA01QP1.
FEATURES Location/Qualifiers
source 1. .1200
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="CS0CAP008YB01"
/tissue_type="THYMUS"
/clone_lib="Homo sapiens THYMUS"
/note="Vector: pCMVSPORT_6; 1st strand cDNA was primed with a NotI-oligo(dT) primer. Five prime end enriched, double-strand cDNA was digested with Not I and cloned into the Not I and EcoRV sites of the pCMVSPORT 6 vector. Library was not normalized."

ORIGIN

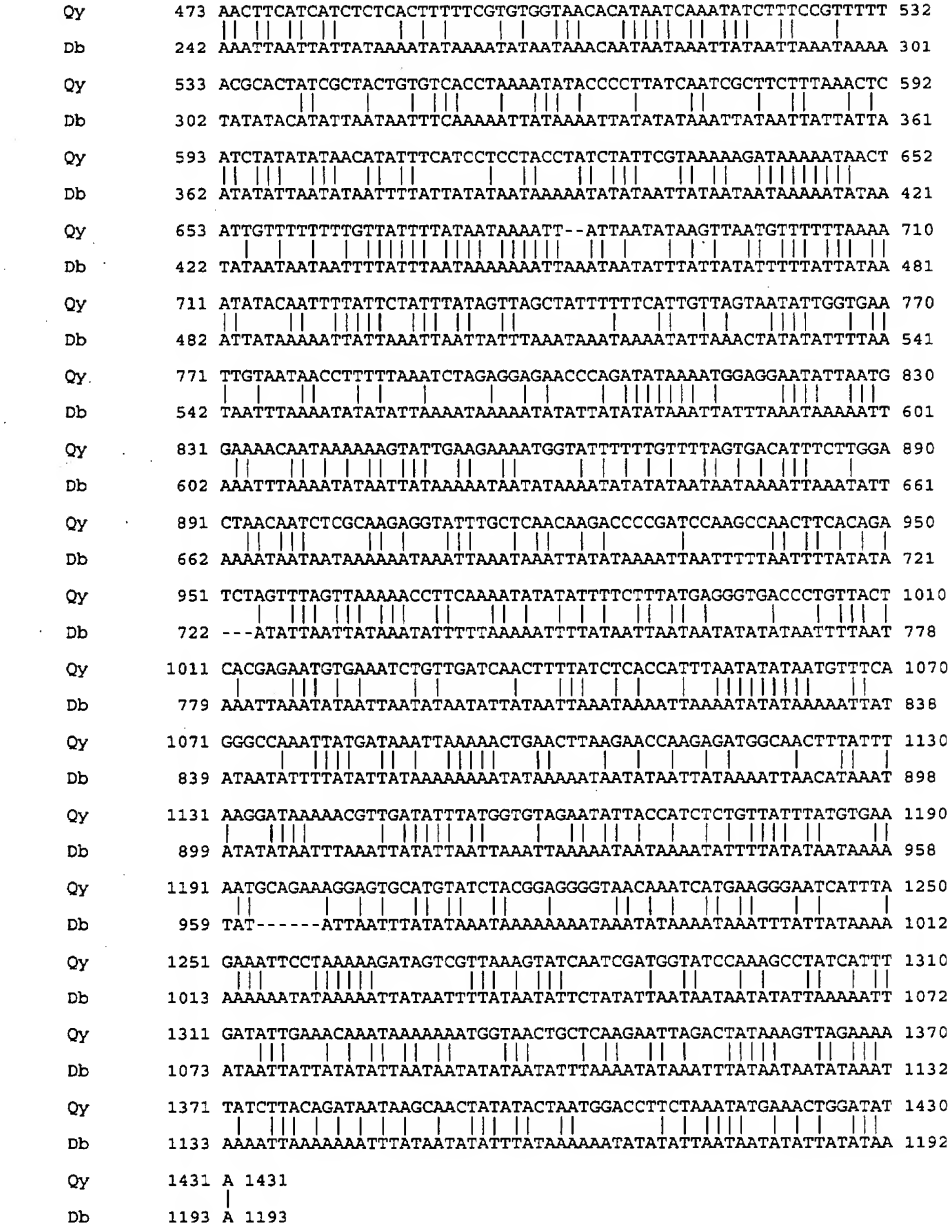
Query Match 5.2%; Score 96.8; DB 13; Length 1200;
Best Local Similarity 31.5%; Pred. No. 8.2e-06;
Matches 253; Conservative 161; Mismatches 382; Indels 6; Gaps 1;





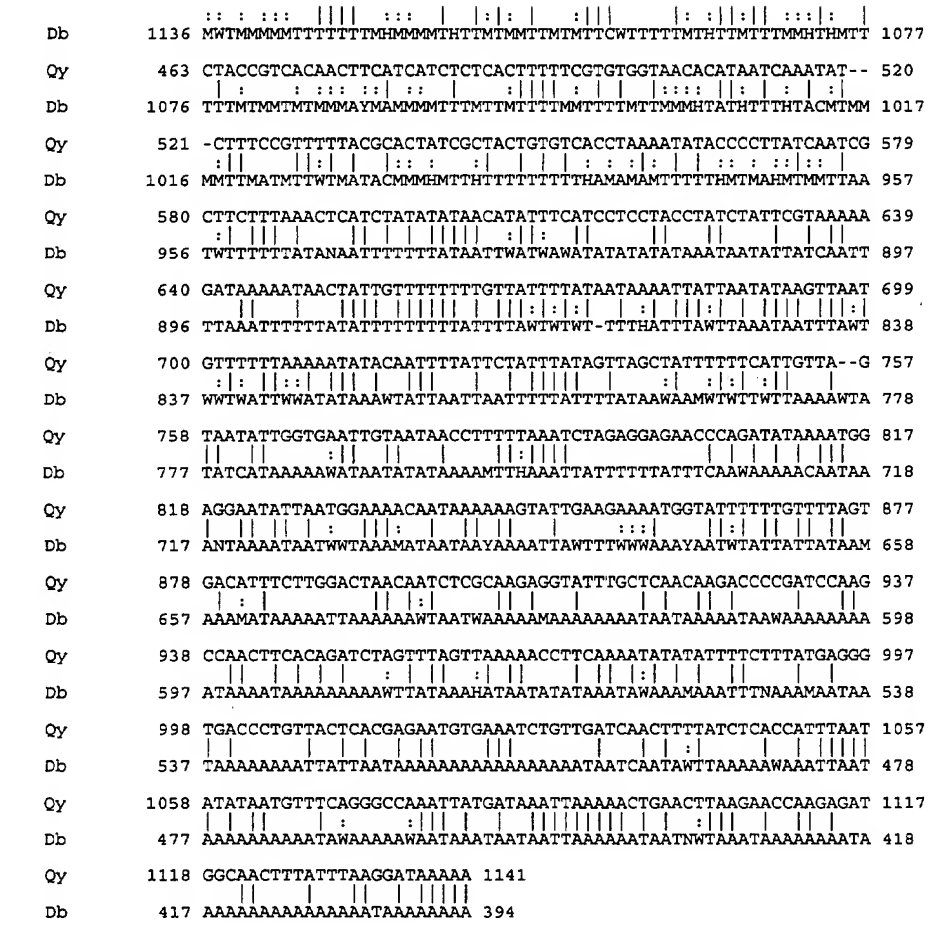
RESULT 3
CC253231
LOCUS CC253231 1277 bp DNA linear GSS 13-MAY-2003
DEFINITION CH261-180N11_RM1.1 CH261 Gallus gallus genomic clone CH261-180N11, genomic survey sequence.
ACCESSION CC253231
VERSION CC253231.1 GI:30589981
KEYWORDS GSS.
SOURCE Gallus gallus (chicken)
ORGANISM Gallus gallus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Archosauria; Aves; Neognathae; Galliformes; Phasianidae; Phasianinae; Gallus.
REFERENCE 1 (bases 1 to 1277)
AUTHORS Kremitzki,C., Higginbotham,J., Wylie,K., Carter,J., McPherson,J., Warren,W., Graves,T., Mardis,E. and Wilson,R.
TITLE Gallus gallus BAC End Reads
JOURNAL Unpublished (2003)
COMMENT Contact: Richard K. Wilson
Genome Sequencing Center
Washington University School of Medicine
Email: submissions@watson.wustl.edu
Insert Length: 182000 Std Error: 0.00
Seq primer: RM1 TACGACTCACTATAGGGAGA
Class: BAC ends
High quality sequence start: 24
High quality sequence stop: 85.
FEATURES Location/Qualifiers
source 1. .1277
/organism="Gallus gallus"
/mol_type="genomic DNA"
/strain="Red Jungle Fowl"
/db_xref="taxon:9031"
/clone="CH261-180N11"
/sex="female"
/cell_line="UCD001, inbred 256"
/clone_lib="CH261"
/note="Vector: pTARBAC2.1; Site_1: EcoRI; Site_2: EcoRI; CH261 Female Chicken library - for library and clone ordering information: http://www.chori.org/bacpac"
ORIGIN

Query Match 5.2%; Score 96.2; DB 28; Length 1277;
Best Local Similarity 46.0%; Pred. No. 9.9e-06;
Matches 442; Conservative 0; Mismatches 508; Indels 11; Gaps 3;



RESULT 4
CNS0039G
LOCUS CNS0039G 1101 bp DNA linear GSS 03-JUN-1999
DEFINITION Drosophila melanogaster genome survey sequence TET3 end of BAC #

```
            BACR08K10 of RPCI-98 library from Drosophila melanogaster (fruit
            fly), genomic survey sequence.
ACCESSION   AL063921
VERSION     AL063921.1  GI:4941778
KEYWORDS    GSS.
SOURCE      Drosophila melanogaster (fruit fly)
  ORGANISM  Drosophila melanogaster
            Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
            Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
            Ephydroidea; Drosophilidae; Drosophila.
REFERENCE   1  (bases 1 to 1101)
  AUTHORS   Genoscope.
  TITLE     Direct Submission
  JOURNAL   Submitted (02-JUN-1999) Genoscope - Centre National de Sequencage :
            BP 191 91006 EVRY cedex - FRANCE (E-mail : seqref@genoscope.cns.fr
            - Web : www.genoscope.cns.fr)
COMMENT     Determination of this BAC-end sequence was carried out as part of a
            collaboration with the Berkeley Drosophila Genome Project (BDGP).
            The BDGP is constructing a physical map of the Drosophila
            melanogaster genome using these BACs. For further information
            please see  http://www.fruitfly.org The BDGP Drosophila
            melanogaster BAC library was prepared by Kazutoyo Osoegawa and
            Aaron Mammoser in Pieter de Jong's laboratory in the Department of
            Cancer Genetics at the Roswell Park Cancer Institute in Buffalo,
            NY. The library is named RPCI-98 and was constructed by partial
            EcoRI digestion of Drosophila DNA provided by the BDGP from the
            isogenic strain y2; cn bw sp, the same strain used for the BDGP's
            P1 and EST libraries. A more detailed description of the library
            and how to order individual BAC clones, the entire library, or
            filters for hybridization from the BACPAC Resource Center can be
            found at http://bacpac.med.buffalo.edu/drosophila_bac.htm.
FEATURES             Location/Qualifiers
     source          1. .1101
                     /organism="Drosophila melanogaster"
                     /mol_type="genomic DNA"
                     /db_xref="taxon:7227"
                     /clone="BACR08K10"
                     /clone_lib="RPCI-98"
                     /note="end : TET3"
ORIGIN

  Query Match             5.1%;  Score 93.8;  DB 29;  Length 1101;
  Best Local Similarity   18.3%;  Pred. No. 2.5e-05;
  Matches  123;  Conservative  287;  Mismatches  261;  Indels    0;  Gaps    0;

Qy         81 AACATAGATTAGGGCATGGAGATTTACCAGACAACTATGAACGTATATACTCACATCACG 140
Db        421 WAAAATAATTWWAAWAWAAAAAATTWWAAAAWAAAAWAWTAWWTTTAWTWAAAAAAAAAA 480

Qy        141 CAATCGGCAATTGATGACATTGGAACTAAATTCAATCAATTTGTTACTAACAAGCAACTA 200
Db        481 AATTWTTTTTTTWTTTAWTTWATAWWTTWWWTTAAAWAAAAAAAAAAAAWAAAWAAAWATA 540

Qy        201 GATTGACAACTAATTCTCAACAAACGTTAATTTAACAACATTCAAGTAACTCCCACCAGC 260
Db        541 AATWTWWTWWTTYTTTWAAWATAAAMCMAAWYYHTYTTYTYHYYTYWTYTMTWHYHTMYTH 600

Qy        261 TCCATCAATGCTTACCGTAAGTAATCATAACTTACTAAAACCTTGTTACATCAAGGTTTT 320
Db        601 AWAHTTWYHWYHTYAMWHWMTWHWTMWAWWHWTTYTAAYYYYYTCMYYYHYMHWHHAHAH 660

Qy        321 TTCTTTTTGTCTTGTTCATGAGTTACCATAACTTTCTATATTATTGACAACTAAATTGAC 380
Db        661 AAAWWTTHTWWTHAYHWATYHYYYYMYCAMMCMCTHTCHHCYYYYHHYTAHHTHTHHWYA 720

Qy        381 AACTCTTCAATTATTTTTCTGTCTACTCAAAGTTTTCTTCATTTGATATAGTCTAATTCC 440
Db        721 HYYMWYYMWAYYWMYCTACTYHYHHHHHYHWAYHTTWYAWAHAMWMWHHAHYAAAAAWAA 780

Qy        441 ACCATCACTTCTTCCACTCTCTCTACCGTCACAACTTCATCATCTCTCACTTTTTCGTGT 500
Db        781 WATTHHYHHTTHYMHHTYMYHYYMYTCCYMCTYHCWHYYHTAYTCWTWTHHWMWTWTHWY 840

Qy        501 GGTAACACATAATCAAATATCTTTCCGTTTTTACGCACTATCGCTACTGTGTCACCTAAA 560
Db        841 HHTWWHHTTTHWAWWHTHTWCWWWWHATTWTWATHCWACMTMHWHHWMHMHHHHMACHAH 900

Qy        561 ATATACCCCTTATCAATCGCTTCTTTAAACTCATCTATATATAACATATTTCATCCTCCT 620
Db        901 HHTHMCMCHHHMHCTCHHHHTMYHMTCHWMWHMHWHWHHWWMATWMTTMTTMMMMCCMMH 960

Qy        621 ACCTATCTATTCGTAAAAAGATAAAAATAACTATTGTTTTTTTTGTTATTTTATAATAAA 680
Db        961 HHCHMYHMMHMYMYCCHYYCTCHTHATTHYHYMCTCYHYCTWHTYWTAYWWAWTAHAMTT 1020

Qy        681 ATTATTAATATAAGTTAATGTTTTTTAAAAATATACAATTTTATTCTATTTATAGTTAGC 740
Db       1021 ATWWWWMHWWAHWATWWWWWWWWATAWWACTCHHWTWYHTHCTWYYHHTYHMWWAWWMAW 1080

Qy        741 TATTTTTTCAT 751
Db       1081 HWHHMYAHYHW 1091


RESULT 5
BX437739/c
LOCUS       BX437739               1200 bp    mRNA    linear   EST 22-MAY-2003
DEFINITION  BX437739 Homo sapiens THYMUS Homo sapiens cDNA clone CS0CAP008YG24
            3-PRIME, mRNA sequence.
ACCESSION   BX437739
VERSION     BX437739.1  GI:31018315
KEYWORDS    EST.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 1200)
  AUTHORS   Li,W.B., Gruber,C., Jessee,J. and Polayes,D.
  TITLE     Full-length cDNA libraries and normalization
  JOURNAL   Unpublished (2001)
COMMENT     Contact: Genoscope
            Genoscope - Centre National de Sequencage
            BP 191 91006 EVRY cedex - France
            Email: seqref@genoscope.cns.fr, Web : www.genoscope.cns.fr
            Library was constructed by Life Technologies, a division of
            Invitrogen. This sequence belongs to sequence cluster 534.r For
            more information about this cluster, see
            http://www.genoscope.cns.fr/
            cgi-bin/cluster.cgi?seq=CS0CAP008BD12NP1&cluster=534.r. Contact :
            Feng Liang Email : fliang@lifetech.com URL :
            http://fulllength.invitrogen.com/ InVitroGen Corporation 1600
            Faraday Avenue Genoscope sequence ID : CS0CAP008BD12NP1.
FEATURES             Location/Qualifiers
     source          1. .1200
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
                     /clone="CS0CAP008YG24"
                     /tissue_type="THYMUS"
                     /clone_lib="Homo sapiens THYMUS"
                     /note="Vector: pCMVSPORT_6; 1st strand cDNA was primed
                     with a NotI-oligo(dT) primer. Five prime end enriched,
                     double-strand cDNA was digested with Not I and cloned into
                     the Not I and EcoRV sites of the pCMVSPORT 6 vector.
                     Library was not normalized."
ORIGIN

  Query Match             4.9%;  Score 91.6;  DB 13;  Length 1200;
  Best Local Similarity   37.4%;  Pred. No. 5.3e-05;
  Matches  301;  Conservative  110;  Mismatches  387;  Indels    6;  Gaps    3;

Qy        343 TTACCATAACTTTCTATATTATTGACAACTAAATTGACAACTCTTCAATTATTTTTCTGT 402
Db       1196 TTTTMMMMHMHWTATTCMTTHMTATHHMMTHMTMTHTHWHTTHMMMWTTTTTTTMMMTTM 1137

Qy        403 CTACTCAAAGTTTTCTTCATTTGATATAGTCTAATTCCACCATCACTTCTTCCACTCTCT 462
```



RESULT 6
BX355654/c
LOCUS BX355654 1201 bp mRNA linear EST 05-MAY-2003
DEFINITION BX355654 Homo sapiens PLACENTA COT 25-NORMALIZED Homo sapiens cDNA clone CS0DI002YE12 3-PRIME, mRNA sequence.
ACCESSION BX355654
VERSION BX355654.1 GI:30371982
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 1201)
AUTHORS Li,W.B., Gruber,C., Jessee,J. and Polayes,D.
TITLE Full-length cDNA libraries and normalization
JOURNAL Unpublished (2001)
COMMENT Contact: Genoscope
Genoscope - Centre National de Sequencage
BP 191 91006 EVRY cedex - France
Email: seqref@genoscope.cns.fr, Web : www.genoscope.cns.fr
Library was constructed by Life Technologies, a division of Invitrogen. Contact : Feng Liang Email : fliang@lifetech.com URL : http://fulllength.invitrogen.com/ InVitroGen Corporation 1600 Faraday Avenue Genoscope sequence ID : CS0DI002BC06NP1.
FEATURES Location/Qualifiers
source 1. .1201
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="CS0DI002YE12"
/tissue_type="PLACENTA COT 25-NORMALIZED"
/clone_lib="Homo sapiens PLACENTA COT 25-NORMALIZED"
/note="1st strand cDNA was primed with a NotI-oligo(dT) primer. Five prime end enriched, double-strand cDNA was digested with Not I and cloned into the Not I and EcoR V sites of the pCMVSPORT 6 vector. Library was normalized."
ORIGIN

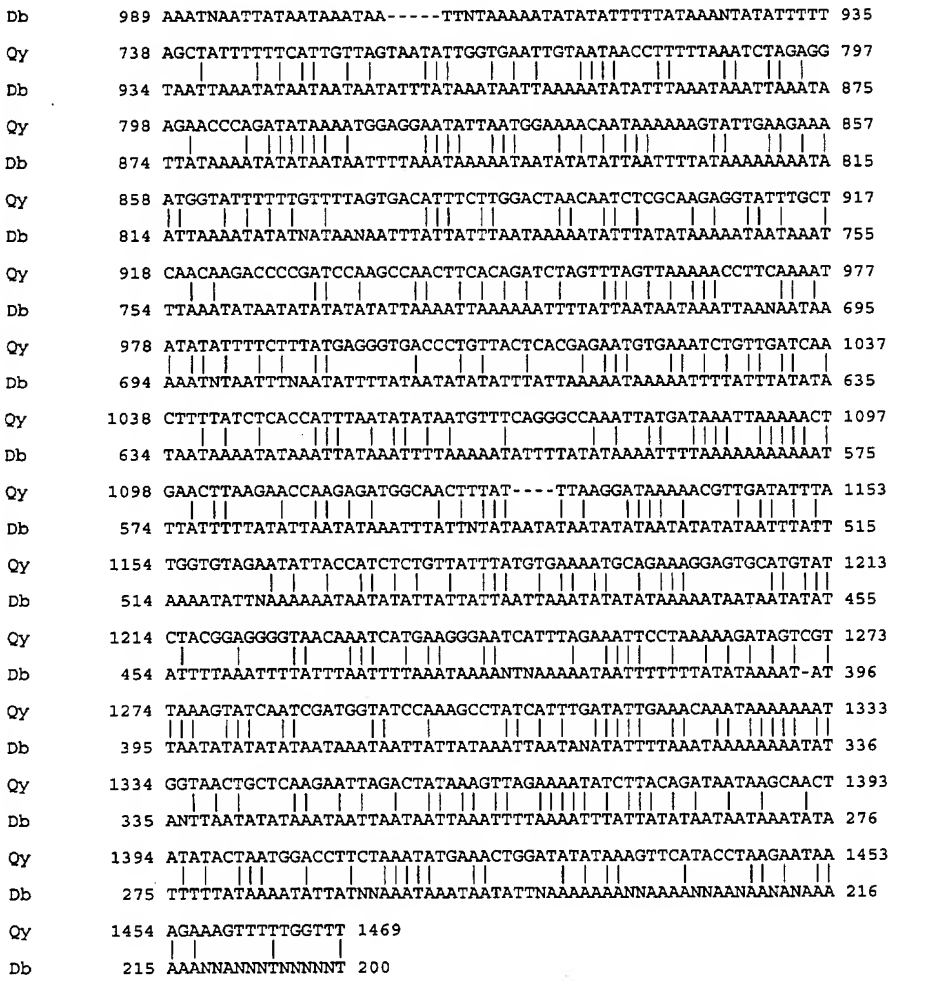
Query Match 4.9%; Score 91.4; DB 13; Length 1201;
Best Local Similarity 31.4%; Pred. No. 5.7e-05;
Matches 217; Conservative 175; Mismatches 290; Indels 10; Gaps 3;

RESULT 7
BX443774/c
LOCUS BX443774 1201 bp mRNA linear EST 15-MAY-2003

```
DEFINITION  BX443774 Homo sapiens T CELLS (JURKAT CELL LINE) Homo sapiens cDNA
            clone CS0DH007YF12 5-PRIME, mRNA sequence.
ACCESSION   BX443774
VERSION     BX443774.1  GI:30772178
KEYWORDS    EST.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 1201)
  AUTHORS   Li,W.B., Gruber,C., Jessee,J. and Polayes,D.
  TITLE     Full-length cDNA libraries and normalization
  JOURNAL   Unpublished (2001)
COMMENT     Contact: Genoscope
            Genoscope - Centre National de Sequencage
            BP 191 91006 EVRY cedex - France
            Email: seqref@genoscope.cns.fr, Web : www.genoscope.cns.fr
            Library was constructed by Life Technologies, a division of
            Invitrogen. This sequence belongs to sequence cluster 7885.f For
            more information about this cluster, see
            http://www.genoscope.cns.fr/
            cgi-bin/cluster.cgi?seq=CS0DH007DC06QP1&cluster=7885.f. Contact :
            Feng Liang Email : fliang@lifetech.com URL :
            http://fulllength.invitrogen.com/ InVitroGen Corporation 1600
            Faraday Avenue Genoscope sequence ID : CS0DH007DC06QP1.
FEATURES             Location/Qualifiers
     source          1. .1201
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
                     /clone="CS0DH007YF12"
                     /tissue_type="T CELLS (JURKAT CELL LINE)"
                     /cell_line="JURKAT CELL LINE"
                     /clone_lib="Homo sapiens T CELLS (JURKAT CELL LINE)"
                     /note="Vector: pCMVSPORT_6; 1st strand cDNA was primed
                     with a NotI-oligo(dT) primer. Five prime end enriched,
                     double-strand cDNA was digested with Not I and cloned into
                     the Not I and EcoRV sites of the pCMVSPORT 6 vector.
                     Library was not normalized."
ORIGIN

  Query Match             4.9%;  Score 91.2;  DB 13;  Length 1201;
  Best Local Similarity   40.9%;  Pred. No. 6.1e-05;
  Matches  255;  Conservative   68;  Mismatches  294;  Indels    7;  Gaps    2;

Qy        587 AAACTCATCTATATATAACATATTTCATCCTCCTACCTATCTATTCGTAAAAAGATAAAA 646
              |||    |   ||:|:||| |::| | :|  |   |  | | : ||   ||||| | ||||
Db       1151 AAAAAWAAAAATWTWTAAAAWWTATTWTWTTAAAATWTTTAWTTTTWAAAAAAAAAAAAA 1092

Qy        647 ATAACTATTGTTTTTTTTGTTATTTTATAATAAAATTATTAATATAAGTTAATGTTTTTT 706
              ||   |:|: ::|:|||| |: |:||||| ||||| || || | |   |:||
Db       1091 ATTTATWTWTWWTWTTTTTTWTTWTTATATTAAAAATAATATTTTTTTTWAAAAAAAAAA 1032

Qy        707 AAAAATATACA------ATTTTATTCTATTTATAGTTAGCTATTTTTTCATTGTTAGTAA 760
              |||:|||||  |       || ||:|: : | ||::  ||    |     ||       :: ||:
Db       1031 AAAWATATATATTATATATATTWTWTWTTATAWWTATATAAAAAAATTTTAAAAWWATAW 972

Qy        761 TATTGGTGAATTGTAATAACCTTTTTAAATCTAGAGGAGAACCCAGATATAAAATGGAGG 820
                :|  | :|:|  |:  |  :::: :|:| :|    | :    : : :: | ||
Db        971 ATWTWTTAWAWTWAAWATAAAWWWWAWAWTWWAWTWTATWTTTTWAWAWWTATATATTTT 912

Qy        821 AATATTAATGGAAAACAATAAAAAAGTATTGAAGAAAATGGTATTTTTTGTTTTAGTGAC 880
              |||| |    |||| || |:||:: |::| || |:||    || ||:| |:|| : |
Db        911 TATATAATAAAAAAAAAAAAWAAWWTTWWTTAAAAWAAAAAAATATTWTTTWTTTTWTAT 852

Qy        881 ATTTCTTGGACTAACAATCTCGCAAGAGGTATTTGCTCAACAAGACCCCGATCCAAGCCA 940
              |: : |      || ||      :     |||    : || :| |      |  |:
Db        851 AWAWATATATAAAAAAAAAAAAAWTTTWTTATAAAAWAAAAWATAAWTAWTTTAAWTTWT 792

Qy        941 ACTTCACAGATCTAGTTTAGTTAAAAACCTTCAAAATATATATTTTCTTTATGAGGGTGA 1000
                | | || ||    |     ||:||:  ||      | | :  |:| ::|||     |
Db        791 TTTAAATATATATATATTATAWAAHTTTTTTTTTTTTTWTATDTAKWTATTTTTTTTT 732
```

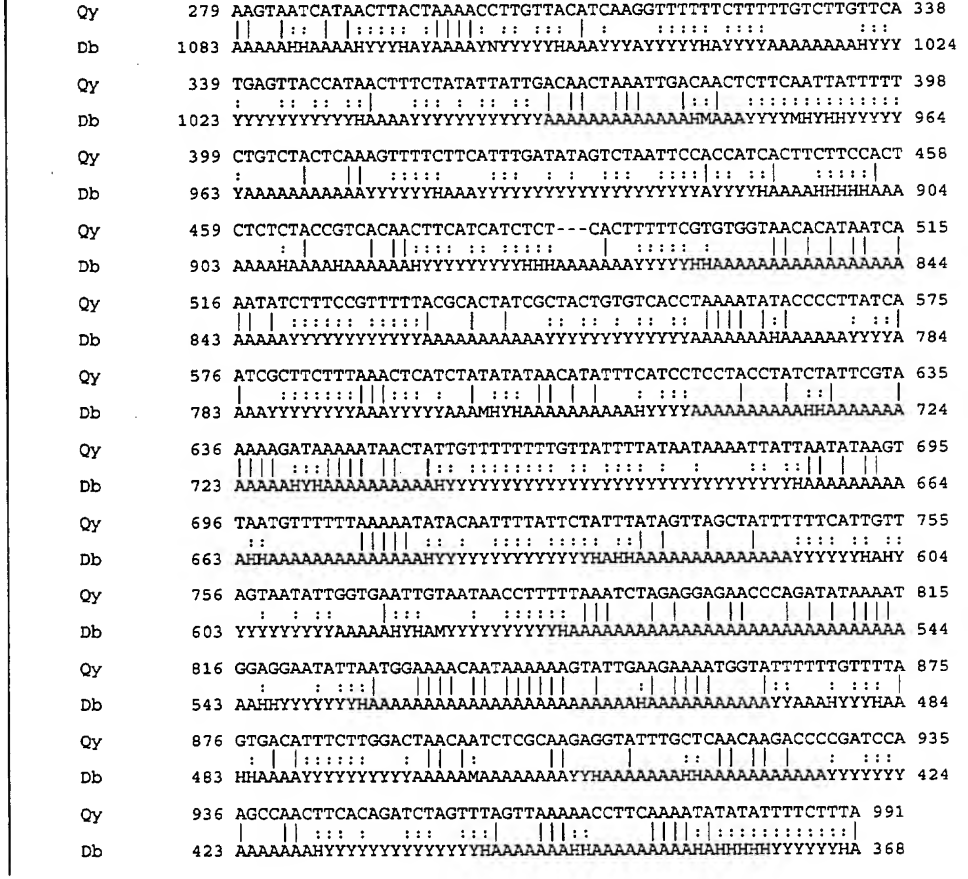
```
Qy       1001 CCCTGTTACTCACGAGAATGTGAAATCTGTTGATCAACTTTTATCTCACCATTTAATATA 1060
                | ||  |       | |    | | || :|    |||| | |     |||  |:|:
Db        731 TTTTTTTTTTTTTTTTTTTTTTTT-TWTTTWTWTTTTTTTTTTTTTTTTTTWTW 673

Qy       1061 TAATGTTTCAGGGCCAAATTATGATAAATTAAAAACTGAACTTAAGAACCAAGAGATGGC 1120
              ||:: : :  ||:  |||  | | |||  |||||   ||   || ||  || | |
Db        672 TAWWAWAKGKGGDWAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 613

Qy       1121 AACTTTATTTAAGGATAAAAACGTTGATATTTATGGTGTAGAATATTACCATCTCTGTTA 1180
              ||    |   ||  | |||||  ||  | ||| |  | |    | ||    |  || | |
Db        612 AAAAAAAAAAAAAAAAAAAAAWTTTTTTTTTTTTTTTTTTTTTTTTTGAAGTTCAGCTT 553

Qy       1181 TTTATGTGAAAATGCAGAAAGGAG 1204
              |||||   | |    | ||| |||
Db        552 TTTATTGAACATGTTATAAAAGAG 529
```

```
RESULT 8
CC262481/c
LOCUS       CC262481               1202 bp    DNA     linear   GSS 13-MAY-2003
DEFINITION  CH261-167M9_Sp6.1 CH261 Gallus gallus genomic clone CH261-167M9,
            genomic survey sequence.
ACCESSION   CC262481
VERSION     CC262481.1  GI:30607397
KEYWORDS    GSS.
SOURCE      Gallus gallus (chicken)
  ORGANISM  Gallus gallus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Archosauria; Aves; Neognathae; Galliformes; Phasianidae;
            Phasianinae; Gallus.
REFERENCE   1  (bases 1 to 1202)
  AUTHORS   Kremitzki,C., Higginbotham,J., Wylie,K., Carter,J., McPherson,J.,
            Warren,W., Graves,T., Mardis,E. and Wilson,R.
  TITLE     Gallus gallus BAC End Reads
  JOURNAL   Unpublished (2003)
COMMENT     Contact: Richard K. Wilson
            Genome Sequencing Center
            Washington University School of Medicine
            Email: submissions@watson.wustl.edu
            Insert Length: 182000   Std Error: 0.00
            Seq primer: Sp6 ATTTAGGTGACACTATAG
            Class: BAC ends
            High quality sequence start: 30
            High quality sequence stop: 105.
FEATURES             Location/Qualifiers
     source          1. .1202
                     /organism="Gallus gallus"
                     /mol_type="genomic DNA"
                     /strain="Red Jungle Fowl"
                     /db_xref="taxon:9031"
                     /clone="CH261-167M9"
                     /sex="female"
                     /cell_line="UCD001, inbred 256"
                     /clone_lib="CH261"
                     /note="Vector: pTARBAC2.1; Site_1: EcoRI; Site_2: EcoRI;
                     CH261 Female Chicken library - for library and clone
                     ordering information: http://www.chori.org/bacpac"
ORIGIN

  Query Match             4.9%;  Score 90.2;  DB 28;  Length 1202;
  Best Local Similarity   44.5%;  Pred. No. 8.7e-05;
  Matches  408;  Conservative    0;  Mismatches  498;  Indels   10;  Gaps    3;

Qy        558 AAAATATACCCCTTATCAATCGCTTCTTTAAACTCATCTATATATAACATATTTCATCCT 617
              ||| | ||       | ||    ||  ||| | || | |||     |  ||  |   |
Db       1109 AAATTTTATATAAAAATAAATAATTAAATAATTTTATATTTATTATTAAATTTAAAAAAT 1050

Qy        618 CCTACCTATCTATTCGTAAAAAGATAAAAATAACTATTGTTTTTTTTGTTATTTTATAAT 677
              ||  ||  || |  |||||| ||| |||| | ||      ||| || |   || |  ||
Db       1049 AATATATAAATAATAATAAAAAAATATAAATTAATAAAAAATTTATTATATNTTAAATAT 990

Qy        678 AAAATTATTAATATAAGTTAATGTTTTTTAAAAATATACAATTTTATTCTATTTATAGTT 737
              |||   ||||  |||| | |     || ||||||||||| | ||||||    | |||  ||
```

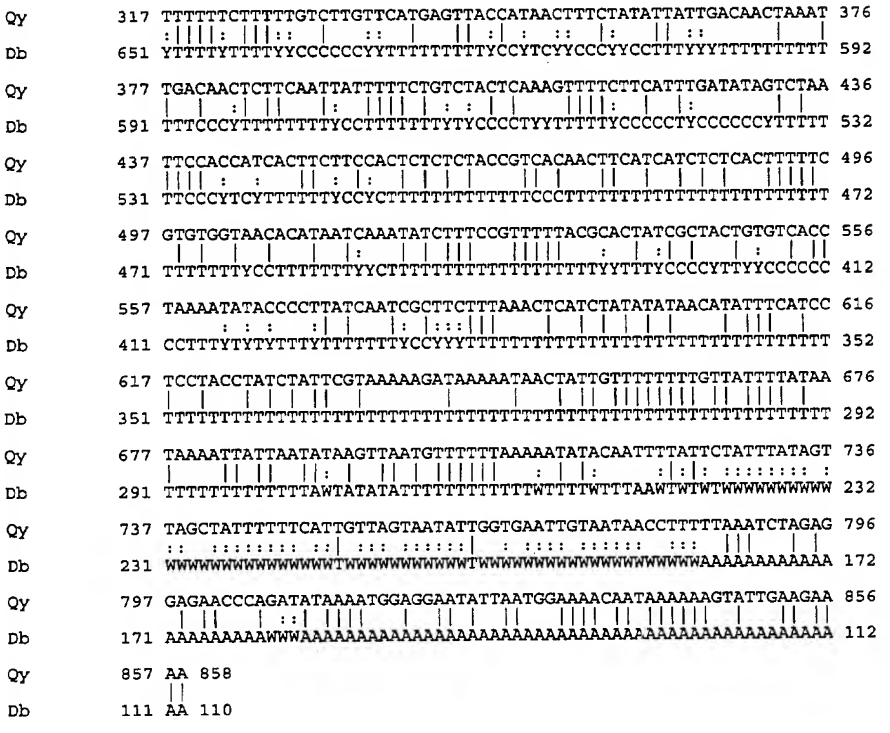


RESULT 9
BX436282/c
LOCUS BX436282 1124 bp mRNA linear EST 15-MAY-2003
DEFINITION BX436282 Homo sapiens THYMUS Homo sapiens cDNA clone CS0CAP001YC01 5-PRIME, mRNA sequence.
ACCESSION BX436282
VERSION BX436282.1 GI:30787521
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 1124)
AUTHORS Li,W.B., Gruber,C., Jessee,J. and Polayes,D.
TITLE Full-length cDNA libraries and normalization
JOURNAL Unpublished (2001)
COMMENT Contact: Genoscope
Genoscope - Centre National de Sequencage
BP 191 91006 EVRY cedex - France
Email: seqref@genoscope.cns.fr, Web : www.genoscope.cns.fr
Library was constructed by Life Technologies, a division of Invitrogen. Contact : Feng Liang Email : fliang@lifetech.com URL : http://fulllength.invitrogen.com/ InVitroGen Corporation 1600 Faraday Avenue Genoscope sequence ID : CS0CAP001AB01QP1.
FEATURES Location/Qualifiers
source 1. .1124
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="CS0CAP001YC01"
/tissue_type="THYMUS"
/clone_lib="Homo sapiens THYMUS"
/note="Vector: pCMVSPORT_6; 1st strand cDNA was primed with a NotI-oligo(dT) primer. Five prime end enriched, double-strand cDNA was digested with Not I and cloned into the Not I and EcoRV sites of the pCMVSPORT 6 vector. Library was not normalized."
ORIGIN

Query Match 4.9%; Score 89.8; DB 13; Length 1124;
Best Local Similarity 21.1%; Pred. No. 0.0001;
Matches 151; Conservative 262; Mismatches 300; Indels 3; Gaps 1;



```
RESULT 10
AL565455
LOCUS       AL565455               1201 bp    mRNA    linear   EST 12-MAY-2003
DEFINITION  AL565455 Homo sapiens FETAL BRAIN Homo sapiens cDNA clone
            CS0DF005YO18 3-PRIME, mRNA sequence.
ACCESSION   AL565455
VERSION     AL565455.2  GI:30549492
KEYWORDS    EST.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 1201)
  AUTHORS   Li,W.B., Gruber,C., Jessee,J. and Polayes,D.
  TITLE     Full-length cDNA libraries and normalization
  JOURNAL   Unpublished (2001)
COMMENT     On Feb 16, 2001 this sequence version replaced gi:12916848.
            Contact: Genoscope
            Genoscope - Centre National de Sequencage
            BP 191 91006 EVRY cedex - France
            Email: seqref@genoscope.cns.fr, Web : www.genoscope.cns.fr
            Library was constructed by Life Technologies, a division of
            Invitrogen. This sequence belongs to sequence cluster 9232.f For
            more information about this cluster, see
            http://www.genoscope.cns.fr/
            cgi-bin/cluster.cgi?seq=CS0DF005BH09NP1&cluster=9232.f. Contact :
            Feng Liang Email : fliang@lifetech.com URL :
            http://fulllength.invitrogen.com/ InVitroGen Corporation 1600
            Faraday Avenue Genoscope sequence ID : CS0DF005BH09NP1.
FEATURES             Location/Qualifiers
     source          1. .1201
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
                     /clone="CS0DF005YO18"
                     /tissue_type="FETAL BRAIN"
                     /dev_stage="fetal"
                     /clone_lib="Homo sapiens FETAL BRAIN"
                     /note="Organ: brain; Vector: pCMVSPORT_6; 1st strand cDNA
                     was primed with a NotI-oligo(dT) primer. Five prime end
                     enriched, double-strand cDNA was digested with Not I and
                     cloned into the Not I and EcoRV sites of the pCMVSPORT 6
                     vector. Library was not normalized."
ORIGIN

  Query Match             4.9%;  Score 89.8;  DB 9;  Length 1201;
  Best Local Similarity   32.4%;  Pred. No. 0.0001;
  Matches  243;  Conservative  135;  Mismatches  372;  Indels    1;  Gaps    1;

Qy        584 TTTAAACTCATCTATATATAACATATTTCATCCTCCTACCTATCTATTCGTAAAAAGATA 643
              | :| |   |: ||:|||||:   : |  ||     :|  :|| :||   :| |:| |:|
Db        435 TAWATAWATAWATAWATATAWATAWATAWATAWATAWATAWATATAWATAWATAWA 494

Qy        644 AAAATAACTATTGTTTTTTTTGTTATTTTATAATAAAATTATTAATATAAGTTAATGTTT 703
              |:|||: ||:   : | | : | : | :|||:  |:|| : ||:    |  ||:   :
Db        495 TAWATAWATAWATAWATATAWATAWATAWATAWATAWATAWATATAWATAWATAWA 554

Qy        704 TTTAAAAATATACAATTTTATTCTATTTATAGTTAGCTATTTTTTCATTGTTAGTAATAT 763
              | :| |:|||:| |: | ||:   : | :   :       : | :   | :   :|  :
Db        555 TAWATAWATAWATAWATATAWATAWATAWATAWATAWATAWATATAWATAWATAWA 614

Qy        764 TGGTGAATTGTAATAACCTTTTTAAATCTAGAGGAGAACCCAGATATAAAATGGAGGAAT 823
              |     |: | :|  :|  | : ||:|| :| |     :|    | |:||| |:   :  |:
Db        615 TAWATAWATAWATAWATATAWATAWATAWATAWATAWATAWATATAWATAWATAWA 674

Qy        824 ATTAATGGAAAACAATAAAAAAGTATTGAAGAAAATGGTATTTTTTGTTTTAGTGACATT 883
              :|     :| |:  | |:|  : | :| |:|     ||: | :   | :|   | |:
Db        675 TAWATAWATAWATAWATATAWATAWATAWATAWATAWATAWATATAWATAWATAWA 734

Qy        884 TCTTGGACTAACAATCTCGCAAGAGGTATTTGCTCAACAAGACCCCGATCCAAGCCAACT 943
              | :   : ||:  |: |   :| | ||: |    |: | |     ||  |     |
Db        735 TAWATAWATAWATAWATATAWATAWATAWATAWATAWATAWATAWATATAWATAWATAWA 794

Qy        944 TCACAGATCTAGTTTAGTTAAAAACCTTCAAAATATATATTTTCTTTATGAGGGTGACCC 1003
              | :   :  : | :|  ||:| |  |  | |:   :   : | : ||| :   :
Db        795 TAWATAWATAWATAWATATAWATAWATAWATAWATAWATAWATATAWATAWATAWA 854

Qy       1004 TGTTACTCACGAGAATGTGAAATCTGTTGATCAACTTTTATCTCACCATTTAATATATAA 1063
              | :   :    | |:    |:|| : |  || :|  : ||: | :  || :|  :   :|
Db        855 TAWATAWATAWATAWATATAWATAWATAWATAWATAWATAWATATAWATAWATAWA 914

Qy       1064 TGTTTCAGGGCCAAATTATGATAAATTAAAAACTGAACTTAAGAACCAAGAGATGGCAAC 1123
              | : | :       |: ||| :  |: ||:| |    |: | :| |:  |  :      |:
Db        915 TAWATAWATAWATAWATATAWATAWATAWATAWATAWATAWATAWATATAWATAWATAWA 974

Qy       1124 TTTATTTAAGGATAAAAACGTTGATATTTATGGTGTAGAATATTACCATCTCTGT-TATT 1182
              | :|| :|   |||:| |  :  |:|| :||     :  |:|| :  || : | : ||:
Db        975 TAWATAWATAWATAWATATAWATAWATAWATAWATAWATAWATATAWATAWATAWA 1034

Qy       1183 TATGTGAAAATGCAGAAAGGAGTGCATGTATCTACGGAGGGGTAACAAATCATGAAGGGA 1242
              ||: | :| |:  | | |  : |  || :|| :|   |    ||:  |   || :|   |
Db       1035 TAWATAWATAWATAWATATAWATAWATAWATAWATAWATAWATATAWATAWATAWA 1094

Qy       1243 ATCATTTAGAAATTCCTAAAAAGATAGTCGTTAAAGTATCAATCGATGGTATCCAAAGCC 1302
                || :| |:||   || |:| |:|    | :|  :|| :||  ||   :  :|
Db       1095 TAWATAWATAWATAWATATAWATAWATAWATAWATAWATAWATATAWATAWATAWA 1154

Qy       1303 TATCATTTGATATTGAAACAAATAAAAAAAT 1333
              ||:   :  |:||  | | |:|||:| |:||
Db       1155 TAWATAWATAWATAWATATAWATAWATAWAT 1185


RESULT 11
BX446296/c
LOCUS       BX446296               1201 bp    mRNA    linear   EST 15-MAY-2003
DEFINITION  BX446296 Homo sapiens NEUROBLASTOMA Homo sapiens cDNA clone
            XCL0BB001ZA03 3-PRIME, mRNA sequence.
ACCESSION   BX446296
VERSION     BX446296.1  GI:30784407
KEYWORDS    EST.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 1201)
  AUTHORS   Li,W.B., Gruber,C., Jessee,J. and Polayes,D.
  TITLE     Full-length cDNA libraries and normalization
  JOURNAL   Unpublished (2001)
COMMENT     Contact: Genoscope
            Genoscope - Centre National de Sequencage
            BP 191 91006 EVRY cedex - France
            Email: seqref@genoscope.cns.fr, Web : www.genoscope.cns.fr
            Library was constructed by Life Technologies, a division of
            Invitrogen. Contact : Feng Liang Email : fliang@lifetech.com URL :
            http://fulllength.invitrogen.com/ InVitroGen Corporation 1600
            Faraday Avenue Genoscope sequence ID : XCL0BB001ZA03FP1.
FEATURES             Location/Qualifiers
     source          1. .1201
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
                     /clone="XCL0BB001ZA03"
                     /tissue_type="NEUROBLASTOMA"
                     /clone_lib="Homo sapiens NEUROBLASTOMA"
                     /note="Vector: pCMVSPORT_6; 1st strand cDNA was primed
                     with a NotI-oligo(dT) primer. Five prime end enriched,
                     double-strand cDNA was digested with Not I and cloned into
                     the Not I and EcoRV sites of the pCMVSPORT 6 vector.
                     Library was not normalized."
ORIGIN

  Query Match             4.8%;  Score 89.6;  DB 13;  Length 1201;
  Best Local Similarity   35.8%;  Pred. No. 0.00011;
  Matches  194;  Conservative   87;  Mismatches  261;  Indels    0;  Gaps    0;
```



```
RESULT 12
AL565455/c
LOCUS       AL565455               1201 bp    mRNA    linear   EST 12-MAY-2003
DEFINITION  AL565455 Homo sapiens FETAL BRAIN Homo sapiens cDNA clone
            CS0DF005YO18 3-PRIME, mRNA sequence.
ACCESSION   AL565455
VERSION     AL565455.2  GI:30549492
KEYWORDS    EST.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 1201)
  AUTHORS   Li,W.B., Gruber,C., Jessee,J. and Polayes,D.
  TITLE     Full-length cDNA libraries and normalization
  JOURNAL   Unpublished (2001)
COMMENT     On Feb 16, 2001 this sequence version replaced gi:12916848.
            Contact: Genoscope
            Genoscope - Centre National de Sequencage
            BP 191 91006 EVRY cedex - France
            Email: seqref@genoscope.cns.fr, Web : www.genoscope.cns.fr
            Library was constructed by Life Technologies, a division of
            Invitrogen. This sequence belongs to sequence cluster 9232.f For
            more information about this cluster, see
            http://www.genoscope.cns.fr/
            cgi-bin/cluster.cgi?seq=CS0DF005BH09NP1&cluster=9232.f. Contact :
            Feng Liang Email : fliang@lifetech.com URL :
            http://fulllength.invitrogen.com/ InVitroGen Corporation 1600
            Faraday Avenue Genoscope sequence ID : CS0DF005BH09NP1.
FEATURES             Location/Qualifiers
     source          1. .1201
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
                     /clone="CS0DF005YO18"
                     /tissue_type="FETAL BRAIN"
                     /dev_stage="fetal"
                     /clone_lib="Homo sapiens FETAL BRAIN"
                     /note="Organ: brain; Vector: pCMVSPORT_6; 1st strand cDNA
                     was primed with a NotI-oligo(dT) primer. Five prime end
                     enriched, double-strand cDNA was digested with Not I and
                     cloned into the Not I and EcoRV sites of the pCMVSPORT 6
                     vector. Library was not normalized."
ORIGIN
```

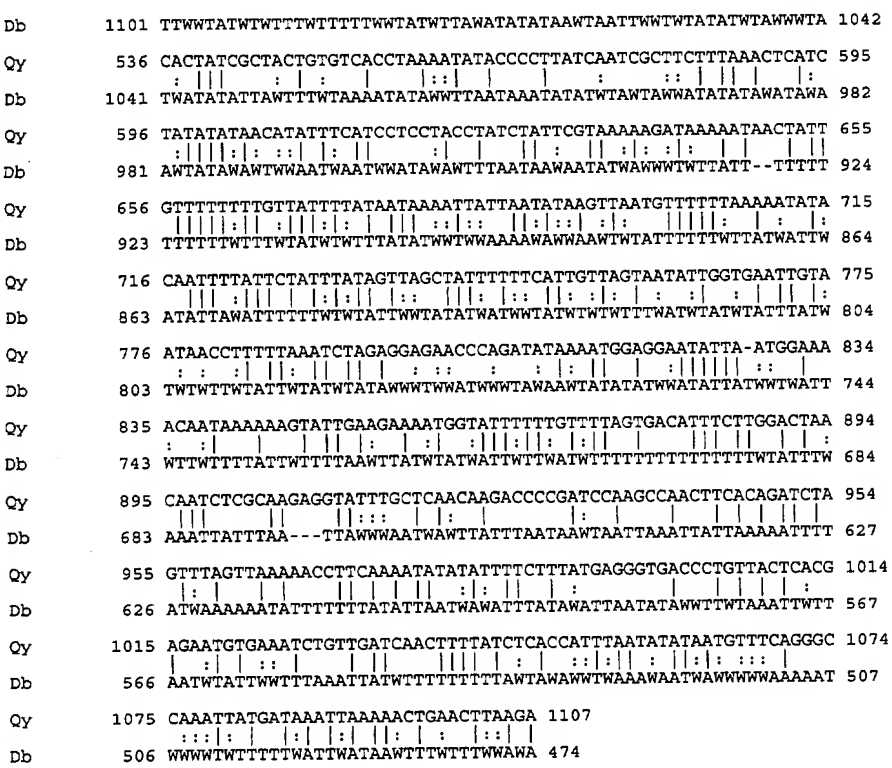


```
  Query Match             4.8%;  Score 88.8;  DB 9;  Length 1201;
  Best Local Similarity   34.2%;  Pred. No. 0.00014;
  Matches  247;  Conservative  123;  Mismatches  350;  Indels    2;  Gaps    2;
```



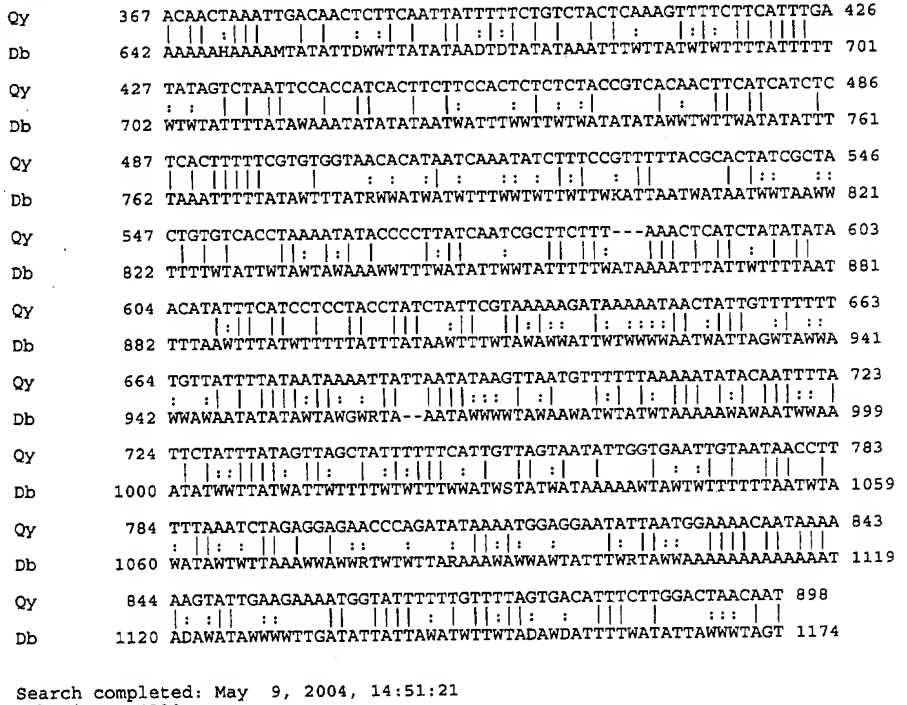
```
RESULT 13
CNS0039G/c
LOCUS       CNS0039G               1101 bp    DNA     linear   GSS 03-JUN-1999
```

```
DEFINITION  Drosophila melanogaster genome survey sequence TET3 end of BAC #
            BACR08K10 of RPCI-98 library from Drosophila melanogaster (fruit
            fly), genomic survey sequence.
ACCESSION   AL063921
VERSION     AL063921.1  GI:4941778
KEYWORDS    GSS.
SOURCE      Drosophila melanogaster (fruit fly)
  ORGANISM  Drosophila melanogaster
            Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
            Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
            Ephydroidea; Drosophilidae; Drosophila.
REFERENCE   1  (bases 1 to 1101)
  AUTHORS   Genoscope.
  TITLE     Direct Submission
  JOURNAL   Submitted (02-JUN-1999) Genoscope - Centre National de Sequencage :
            BP 191 91006 EVRY cedex - FRANCE (E-mail : seqref@genoscope.cns.fr
            - Web : www.genoscope.cns.fr)
COMMENT     Determination of this BAC-end sequence was carried out as part of a
            collaboration with the Berkeley Drosophila Genome Project (BDGP).
            The BDGP is constructing a physical map of the Drosophila
            melanogaster genome using these BACs. For further information
            please see  http://www.fruitfly.org The BDGP Drosophila
            melanogaster BAC library was prepared by Kazutoyo Osoegawa and
            Aaron Mammoser in Pieter de Jong's laboratory in the Department of
            Cancer Genetics at the Roswell Park Cancer Institute in Buffalo,
            NY. The library is named RPCI-98 and was constructed by partial
            EcoRI digestion of Drosophila DNA provided by the BDGP from the
            isogenic strain y2; cn bw sp, the same strain used for the BDGP's
            P1 and EST libraries. A more detailed description of the library
            and how to order individual BAC clones, the entire library, or
            filters for hybridization from the BACPAC Resource Center can be
            found at http://bacpac.med.buffalo.edu/drosophila_bac.htm.
FEATURES             Location/Qualifiers
     source          1. .1101
                     /organism="Drosophila melanogaster"
                     /mol_type="genomic DNA"
                     /db_xref="taxon:7227"
                     /clone="BACR08K10"
                     /clone_lib="RPCI-98"
                     /note="end : TET3"
ORIGIN

  Query Match             4.8%;  Score 88.4;  DB 29;  Length 1101;
  Best Local Similarity   20.3%;  Pred. No. 0.00017;
  Matches  146;  Conservative  295;  Mismatches  269;  Indels    8;  Gaps    2;

Qy        826 TAATGGAAAACAATAAAAAAGTATTGAAGAAAATGGTATTTTTTGTTTTAGTGACATTTC 885
              :|:  |:: : ::|: :::: | :: :::::|:::::| : : : ::: ||| : ||::
Db       1100 KARRWGDDTWDRDTRKDDWDWTKWWTWWKDRADDRRWAGDADRWAWDDGAGTWWTATWWW 1041

Qy        886 TTGGACTAACAATCTCGCAAGAGGTATTTGCTCAACAAGACCCCGATCCAAGCCAACTTC 945
              ::  : |:: ::: :    |  |: :  :|    |: |: |    ||   :::   : :
Db       1040 WWWWWATWDTWWDKWWWWATAAKTDTAWTWWRTAWRADWAGRDRGAGKRDRDAATDADGA 981

Qy        946 ACAGATCTAGTTTAGTTAAAAACCTTCAAAATATATATTTTCTTTATGAGGGTGACCCTG 1005
                ::   :::::::::: :::   :   ||:|  :|::: :::::: : :||   :
Db        980 GRRDGGRKRKDKKDRKDGDDDKKGGKKKKAAKAAKWATKWWDDWDWDKDWKWDGAKDRKA 921

Qy       1006 TTACTCACGAGAATGTGAAATCTGTTGATCAACTTTTATCTCACCATTTAATATATAATG 1065
              :::     :::: :|::|::: :||::::  : ::::::  :    : |:|: ||:::::|
Db        920 DDDDGAGDKDDDGKGKDADDDTDGTKDDDDKDKWDDWDKAKGTWGDATWAWAATDWWWWG 861

Qy       1066 TTTCAGGGCCAAATTATGATAAATTAAAAACTGAACTTAAGAACCAAGAGATGGCAACTT 1125
              : :       |:::|:: | :::: ::::: : :  ::|: :    |:::::|: |
Db        860 WAD-------ADWWTWDAAADDWWADDRWDAWAWKWDDAWAWGARTADRRDWGDRAGKRG 808

Qy       1126 TATTTAAGGATAAAAACGTTGATATTTATGGTGTAGAATATTACCATCTCTGTTATTTAT 1185
              | : ::::: :: ::  :::::| |: | | | :: ::::: :| | :  ::|::::
Db        807 GARKRRDRKRADDKRDAADDRDDAATWTTWTTTTTRDTDDWKWKTDTWTRWAADRTWDRD 748

Qy       1186 GTGAAAATGCAGAAAGGAGTGCATGTATCTACGGAGGGGTAACAAATCATGAAGGGAATC 1245
              ::::::: | ||: :::  :: ::: :| ::  :|::   :: ::: :  :||:|  |
Db        747 DDDDRDRAGTAGRKWRRTWKRRWKRRDTRWDDADADDTARDDRRRRGDDGADAGKGKKTG 688

Qy       1246 ATTTAGAAATTCCTAAAAAGATAGTCGTTAAAGTATCAATCGATGGTATCCAAAGCCTAT 1305
              ::  ::::||:  |:|::|:| : |  :|: ::::: ::  :  :: :   |:|  ::|
Db        687 RKRRRRDRATWDRTDAWWADAAWWTTTDTDTDDWDKRDRRRKGARRRRRTTARAAWDWWT 628

Qy       1306 CATTTGATATTGAAACAAATAAAAAAATGGTAACTGCTCAAGAATTAGACTATAAAGTTA 1365
              :: |::::: :| ::: :||: : : ::|: :  : :| :|  ::|  | |::::|
Db        627 WKAWDWAKWDWKTRADRWDRWAADTWTDARKADRDWAKARAWRARRDRARAARADRRWTT 568

Qy       1366 GAAAATATCTTACAGATAATAAGCAACTATATACTAATGGACCTTCTAAATATGAAACTG 1425
              :    ||| ||: |:| :: ::  |  |||:|  |: |     || |   |:|  || :
Db        567 KGKTTTATWTTWAARAAWWAWWAWATTTATWTT-TWTTTWTTTTTTTTTTTWTTTAAWWW 509

Qy       1426 GATATATAAAGTTCATACCTAAGAATAAAGAAAGTTTTTGGTTTGATTTTTTCCCTGAAC 1485
               |::|||:|| |  |:|   || :|         ||||:  |   ::| :|:   |
Db        508 AAWWTATWAAWTAAAWAAAAAAAWAATTTTTTTTTTTTTWAWTAAAWWTAWTWTTTTWTTT 449

Qy       1486 CAGAATTTACTCAATCTAAATATCTTATGATATATAAAGATAATGAAACGCTTGACTC 1543
              : |||||  |  :| ::|||   || :   |:|| ||| | ::: : :    |: | |
Db        448 TWWAATTTTTTWTWTTWWAATTATTTTWTTTWTAAAAAAAAWWWWWTWTATATKCCCC 391


RESULT 14
CNS00EVL/c
LOCUS       CNS00EVL                1101 bp    DNA     linear   GSS 04-JUN-1999
DEFINITION  Drosophila melanogaster genome survey sequence T7 end of BAC:
            BACR29B23 of RPCI-98 library from Drosophila melanogaster (fruit
            fly), genomic survey sequence.
ACCESSION   AL069706
VERSION     AL069706.1  GI:4949849
KEYWORDS    GSS.
SOURCE      Drosophila melanogaster (fruit fly)
  ORGANISM  Drosophila melanogaster
            Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
            Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
            Ephydroidea; Drosophilidae; Drosophila.
REFERENCE   1  (bases 1 to 1101)
  AUTHORS   Genoscope.
  TITLE     Direct Submission
  JOURNAL   Submitted (02-JUN-1999) Genoscope - Centre National de Sequencage :
            BP 191 91006 EVRY cedex - FRANCE (E-mail : seqref@genoscope.cns.fr
            - Web : www.genoscope.cns.fr)
COMMENT     Determination of this BAC-end sequence was carried out as part of a
            collaboration with the Berkeley Drosophila Genome Project (BDGP).
            The BDGP is constructing a physical map of the Drosophila
            melanogaster genome using these BACs. For further information
            please see  http://www.fruitfly.org The BDGP Drosophila
            melanogaster BAC library was prepared by Kazutoyo Osoegawa and
            Aaron Mammoser in Pieter de Jong's laboratory in the Department of
            Cancer Genetics at the Roswell Park Cancer Institute in Buffalo,
            NY. The library is named RPCI-98 and was constructed by partial
            EcoRI digestion of Drosophila DNA provided by the BDGP from the
            isogenic strain y2; cn bw sp, the same strain used for the BDGP's
            P1 and EST libraries. A more detailed description of the library
            and how to order individual BAC clones, the entire library, or
            filters for hybridization from the BACPAC Resource Center can be
            found at http://bacpac.med.buffalo.edu/drosophila_bac.htm.
FEATURES             Location/Qualifiers
     source          1. .1101
                     /organism="Drosophila melanogaster"
                     /mol_type="genomic DNA"
                     /db_xref="taxon:7227"
                     /clone="BACR29B23"
                     /clone_lib="RPCI-98"
                     /note="end : T7"
ORIGIN

  Query Match             4.7%;  Score 87.8;  DB 29;  Length 1101;
  Best Local Similarity   34.9%;  Pred. No. 0.00021;
  Matches  221;  Conservative  122;  Mismatches  284;  Indels    6;  Gaps    3;

Qy        476 TTCATCATCTCTCACTTTTTCGTGTGGTAACACATAATCAAATATCTTTCCGTTTTTACG 535
              || :|  : : | : ||||:    :   :| | ||||: || |:: :|    :| :::
```



RESULT 15
AL536104
LOCUS AL536104 1201 bp mRNA linear EST 31-MAY-2003
DEFINITION AL536104 Homo sapiens FETAL BRAIN Homo sapiens cDNA clone CS0DF022YC18 5-PRIME, mRNA sequence.
ACCESSION AL536104
VERSION AL536104.2 GI:31260974
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 1201)
AUTHORS Li,W.B., Gruber,C., Jessee,J. and Polayes,D.
TITLE Full-length cDNA libraries and normalization
JOURNAL Unpublished (2001)
COMMENT On Feb 13, 2001 this sequence version replaced gi:12799597.
Contact: Genoscope
Genoscope - Centre National de Sequencage
BP 191 91006 EVRY cedex - France
Email: seqref@genoscope.cns.fr, Web : www.genoscope.cns.fr
Library was constructed by Life Technologies, a division of Invitrogen. Contact : Feng Liang Email : fliang@lifetech.com URL : http://fulllength.invitrogen.com/ InVitroGen Corporation 1600 Faraday Avenue Genoscope sequence ID : CS0DF022BB09QP1.
FEATURES Location/Qualifiers
source 1. .1201
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="CS0DF022YC18"
/tissue_type="FETAL BRAIN"
/dev_stage="fetal"
/clone_lib="Homo sapiens FETAL BRAIN"
/note="Organ: brain; Vector: pCMVSPORT_6; 1st strand cDNA was primed with a NotI-oligo(dT) primer. Five prime end enriched, double-strand cDNA was digested with Not I and cloned into the Not I and EcoRV sites of the pCMVSPORT 6 vector. Library was not normalized."
ORIGIN

Query Match 4.7%; Score 87.8; DB 9; Length 1201;
Best Local Similarity 37.2%; Pred. No. 0.00021;
Matches 199; Conservative 96; Mismatches 235; Indels 5; Gaps 2;



Search completed: May 9, 2004, 14:51:21
Job time : 4339 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: May 8, 2004, 10:40:02 ; Search time 7000 Seconds
(without alignments)
11461.137 Million cell updates/sec

Title: US-10-625-221-12
Perfect score: 1851
Sequence: 1 ccatcacgcatcactcatgt.....ttagcaactattttatcgtc 1851

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 3470272 seqs, 21671516995 residues

Total number of hits satisfying chosen parameters: 6940544

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : GenEmbl:*
1: gb_ba:*
2: gb_htg:*
3: gb_in:*
4: gb_om:*
5: gb_ov:*
6: gb_pat:*
7: gb_ph:*
8: gb_pl:*
9: gb_pr:*
10: gb_ro:*
11: gb_sts:*
12: gb_sy:*
13: gb_un:*
14: gb_vi:*
15: em_ba:*
16: em_fun:*
17: em_hum:*
18: em_in:*
19: em_mu:*
20: em_om:*
21: em_or:*
22: em_ov:*
23: em_pat:*
24: em_ph:*
25: em_pl:*
26: em_ro:*
27: em_sts:*
28: em_un:*
29: em_vi:*
30: em_htg_hum:*
31: em_htg_inv:*
32: em_htg_other:*
33: em_htg_mus:*
34: em_htg_pln:*
35: em_htg_rod:*
36: em_htg_mam:*
37: em_htg_vrt:*
38: em_sy:*
39: em_htgo_hum:*
40: em_htgo_mus:*
41: em_htgo_other:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	1851	100.0	1851	6	AR408359	AR408359 Sequence
2	1851	100.0	1851	6	BD063209	BD063209 Mutants o
3	1835.4	99.2	1837	6	AR089274	AR089274 Sequence
4	1835.4	99.2	1837	6	AR093474	AR093474 Sequence
5	1832.2	99.0	1837	6	AR211577	AR211577 Sequence
6	1832.2	99.0	1837	6	BD223209	BD223209 Bacterial
7	1830.6	98.9	1837	6	AX800046	AX800046 Sequence
8	1668.4	90.1	4621	7	SPU40453	U40453 Streptococc
9	1406.6	76.0	57506	1	AE014161	AE014161 Streptoco
c 10	1406.6	76.0	310850	1	AP005142	AP005142 Streptoco
c 11	1225.4	66.2	11900	1	AE009982	AE009982 Streptoco
12	892.8	48.2	1031	1	SPSPEA	X03929 Streptococc
13	754.4	40.8	756	6	AX110260	AX110260 Sequence
14	706.4	38.2	708	1	SP156SPEA	X61560 S.pyogenes
15	706.4	38.2	708	1	SP165SPEA	X61556 S.pyogenes
16	706.4	38.2	708	1	SP167SPEA	X61559 S.pyogenes
17	706.4	38.2	708	1	SP327SPEA	X61555 S.pyogenes
18	706.4	38.2	708	1	SP493SPEA	X61557 S.pyogenes
19	706.4	38.2	708	1	SP494SPEA	X61558 S.pyogenes
20	704.8	38.1	708	1	SP158SPEA	X61568 S.pyogenes
21	704.8	38.1	708	1	SP250SPEA	X61561 S.pyogenes
22	704.8	38.1	708	1	SP251SPEA	X61562 S.pyogenes
23	704.8	38.1	708	1	SP256SPEA	X61563 S.pyogenes
24	704.8	38.1	708	1	SP285SPEA	X61564 S.pyogenes
25	704.8	38.1	708	1	SP480SPEA	X61565 S.pyogenes
26	704.8	38.1	708	1	SP485SPEA	X61569 S.pyogenes
27	704.8	38.1	708	1	SP491SPEA	X61570 S.pyogenes
28	704.8	38.1	708	1	SP492SPEA	X61566 S.pyogenes
29	704.8	38.1	708	1	SP495SPEA	X61571 S.pyogenes
30	704.8	38.1	708	1	SP496SPEA	X61567 S.pyogenes
31	704.8	38.1	708	1	SP624SPEA	X61572 S.pyogenes
32	663.8	35.9	667	1	AF055698	AF055698 Streptoco
33	658.4	35.6	660	1	SP500SPEA	X61554 S.pyogenes
34	656.8	35.5	1419	6	AX800054	AX800054 Sequence
35	599.2	32.4	708	1	SP262SPEA	X61573 S.pyogenes
36	566.2	30.6	667	1	AF029051	AF029051 Streptoco
37	566.2	30.6	667	1	AY049745	AY049745 Streptoco
38	310.6	16.8	50260	1	AE014155	AE014155 Streptoco
39	310.6	16.8	311650	1	AP005144	AP005144 Streptoco
40	299	16.2	9220	1	AE010048	AE010048 Streptoco
c 41	295.2	15.9	310950	1	AP005143	AP005143 Streptoco
42	288.8	15.6	9911	1	AE010052	AE010052 Streptoco
43	278	15.0	1497	7	BTU22342	U22342 Bacteriopha
44	184.2	10.0	1712	1	STAENTB	M11118 S.aureus en
45	184.2	10.0	1712	6	AR211572	AR211572 Sequence

ALIGNMENTS

RESULT 1
AR408359
LOCUS AR408359 1851 bp DNA linear PAT 18-DEC-2003
DEFINITION Sequence 12 from patent US 6632441.
ACCESSION AR408359
VERSION AR408359.1 GI:40158507
KEYWORDS .
SOURCE Unknown.
ORGANISM Unknown.
Unclassified.
REFERENCE 1 (bases 1 to 1851)
AUTHORS Schlievert,P.M., Roggiani,M., Stoehr,J. and Ohlendorf,D.
TITLE Mutants of streptococcal toxin a and methods of use
JOURNAL Patent: US 6632441-A 12 14-OCT-2003;
FEATURES Location/Qualifiers

applicaB

```
     source          1. .1851
                     /organism="unknown"
                     /mol_type="genomic DNA"
ORIGIN

  Query Match             100.0%;  Score 1851;  DB 6;  Length 1851;
  Best Local Similarity   100.0%;  Pred. No. 3.2e-256;
  Matches 1851;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 CCATCACGCATCACTCATGTTTGACAGCTTATCATCGATAAGCTTACTTTTCGAATCAGG 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 CCATCACGCATCACTCATGTTTGACAGCTTATCATCGATAAGCTTACTTTTCGAATCAGG 60

Qy         61 TCTATCCTTGAAACAGGTGCAACATAGATTAGGGCATGGAGATTTACCAGACAACTATGA 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 TCTATCCTTGAAACAGGTGCAACATAGATTAGGGCATGGAGATTTACCAGACAACTATGA 120

Qy        121 ACGTATATACTCACATCACGCAATCGGCAATTGATGACATTGGAACTAAATTCAATCAAT 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 ACGTATATACTCACATCACGCAATCGGCAATTGATGACATTGGAACTAAATTCAATCAAT 180

Qy        181 TTGTTACTAACAAGCAACTAGATTGACAACTAATTCTCAACAAACGTTAATTTAACAACA 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 TTGTTACTAACAAGCAACTAGATTGACAACTAATTCTCAACAAACGTTAATTTAACAACA 240

Qy        241 TTCAAGTAACTCCCACCAGCTCCATCAATGCTTACCGTAAGTAATCATAACTTACTAAAA 300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 TTCAAGTAACTCCCACCAGCTCCATCAATGCTTACCGTAAGTAATCATAACTTACTAAAA 300

Qy        301 CCTTGTTACATCAAGGTTTTTTCTTTTTGTCTTGTTCATGAGTTACCATAACTTTCTATA 360
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        301 CCTTGTTACATCAAGGTTTTTTCTTTTTGTCTTGTTCATGAGTTACCATAACTTTCTATA 360

Qy        361 TTATTGACAACTAAATTGACAACTCTTCAATTATTTTTCTGTCTACTCAAAGTTTTCTTC 420
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        361 TTATTGACAACTAAATTGACAACTCTTCAATTATTTTTCTGTCTACTCAAAGTTTTCTTC 420

Qy        421 ATTTGATATAGTCTAATTCCACCATCACTTCTTCCACTCTCTCTACCGTCACAACTTCAT 480
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        421 ATTTGATATAGTCTAATTCCACCATCACTTCTTCCACTCTCTCTACCGTCACAACTTCAT 480

Qy        481 CATCTCTCACTTTTTCGTGTGGTAACACATAATCAAATATCTTTCCGTTTTTACGCACTA 540
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        481 CATCTCTCACTTTTTCGTGTGGTAACACATAATCAAATATCTTTCCGTTTTTACGCACTA 540

Qy        541 TCGCTACTGTGTCACCTAAAATATACCCCTTATCAATCGCTTCTTTAAACTCATCTATAT 600
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        541 TCGCTACTGTGTCACCTAAAATATACCCCTTATCAATCGCTTCTTTAAACTCATCTATAT 600

Qy        601 ATAACATATTTCATCCTCCTACCTATCTATTCGTAAAAAGATAAAAATAACTATTGTTTT 660
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        601 ATAACATATTTCATCCTCCTACCTATCTATTCGTAAAAAGATAAAAATAACTATTGTTTT 660

Qy        661 TTTTGTTATTTTATAATAAAATTATTAATATAAGTTAATGTTTTTTAAAAATATACAATT 720
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        661 TTTTGTTATTTTATAATAAAATTATTAATATAAGTTAATGTTTTTTAAAAATATACAATT 720

Qy        721 TTATTCTATTTATAGTTAGCTATTTTTTCATTGTTAGTAATATTGGTGAATTGTAATAAC 780
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        721 TTATTCTATTTATAGTTAGCTATTTTTTCATTGTTAGTAATATTGGTGAATTGTAATAAC 780

Qy        781 CTTTTTAAATCTAGAGGAGAACCCAGATATAAAATGGAGGAATATTAATGGAAAACAATA 840
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        781 CTTTTTAAATCTAGAGGAGAACCCAGATATAAAATGGAGGAATATTAATGGAAAACAATA 840

Qy        841 AAAAAGTATTGAAGAAAATGGTATTTTTTGTTTTAGTGACATTTCTTGGACTAACAATCT 900
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        841 AAAAAGTATTGAAGAAAATGGTATTTTTTGTTTTAGTGACATTTCTTGGACTAACAATCT 900

Qy        901 CGCAAGAGGTATTTGCTCAACAAGACCCCGATCCAAGCCAACTTCACAGATCTAGTTTAG 960
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        901 CGCAAGAGGTATTTGCTCAACAAGACCCCGATCCAAGCCAACTTCACAGATCTAGTTTAG 960

Qy        961 TTAAAAACCTTCAAAATATATATTTTCTTTATGAGGGTGACCCTGTTACTCACGAGAATG 1020
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        961 TTAAAAACCTTCAAAATATATATTTTCTTTATGAGGGTGACCCTGTTACTCACGAGAATG 1020

Qy       1021 TGAAATCTGTTGATCAACTTTTATCTCACCATTTAATATATAATGTTTCAGGGCCAAATT 1080
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1021 TGAAATCTGTTGATCAACTTTTATCTCACCATTTAATATATAATGTTTCAGGGCCAAATT 1080

Qy       1081 ATGATAAATTAAAAACTGAACTTAAGAACCAAGAGATGGCAACTTTATTTAAGGATAAAA 1140
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1081 ATGATAAATTAAAAACTGAACTTAAGAACCAAGAGATGGCAACTTTATTTAAGGATAAAA 1140

Qy       1141 ACGTTGATATTTATGGTGTAGAATATTACCATCTCTGTTATTTATGTGAAAATGCAGAAA 1200
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1141 ACGTTGATATTTATGGTGTAGAATATTACCATCTCTGTTATTTATGTGAAAATGCAGAAA 1200

Qy       1201 GGAGTGCATGTATCTACGGAGGGGTAACAAATCATGAAGGGAATCATTTAGAAATTCCTA 1260
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1201 GGAGTGCATGTATCTACGGAGGGGTAACAAATCATGAAGGGAATCATTTAGAAATTCCTA 1260

Qy       1261 AAAAGATAGTCGTTAAAGTATCAATCGATGGTATCCAAAGCCTATCATTTGATATTGAAA 1320
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1261 AAAAGATAGTCGTTAAAGTATCAATCGATGGTATCCAAAGCCTATCATTTGATATTGAAA 1320

Qy       1321 CAAATAAAAAAAATGGTAACTGCTCAAGAATTAGACTATAAAGTTAGAAAATATCTTACAG 1380
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1321 CAAATAAAAAAAATGGTAACTGCTCAAGAATTAGACTATAAAGTTAGAAAATATCTTACAG 1380

Qy       1381 ATAATAAGCAACTATATACTAATGGACCTTCTAAATATGAAACTGGATATATAAAGTTCA 1440
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1381 ATAATAAGCAACTATATACTAATGGACCTTCTAAATATGAAACTGGATATATAAAGTTCA 1440

Qy       1441 TACCTAAGAATAAAGAAAGTTTTTGGTTTGATTTTTTCCCTGAACCAGAATTTACTCAAT 1500
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1441 TACCTAAGAATAAAGAAAGTTTTTGGTTTGATTTTTTCCCTGAACCAGAATTTACTCAAT 1500

Qy       1501 CTAAATATCTTATGATATATAAAGATAATGAAACGCTTGACTCAAACACAAGCCAAATTG 1560
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1501 CTAAATATCTTATGATATATAAAGATAATGAAACGCTTGACTCAAACACAAGCCAAATTG 1560

Qy       1561 AAGTCTACCTAACAACCAAGTAACTTTTTGCTTTTGGCAACCTTACCTACTGCTGGATTT 1620
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1561 AAGTCTACCTAACAACCAAGTAACTTTTTGCTTTTGGCAACCTTACCTACTGCTGGATTT 1620

Qy       1621 AGAAATTTTATTGCAATTCTTTTATTAATGTAAAAACCGCTCATTTGATGAGCGGTTTTG 1680
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1621 AGAAATTTTATTGCAATTCTTTTATTAATGTAAAAACCGCTCATTTGATGAGCGGTTTTG 1680

Qy       1681 TCTTATCTAAAGGAGCTTTACCTCCTAATGCTGCAAAATTTTAAATGTTGGATTTTTGTA 1740
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1681 TCTTATCTAAAGGAGCTTTACCTCCTAATGCTGCAAAATTTTAAATGTTGGATTTTTGTA 1740

Qy       1741 TTTGTCTATTGTATTTGATGGGTAATCCCATTTTTCGACAGACATCGTCGTGCCACCTCT 1800
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1741 TTTGTCTATTGTATTTGATGGGTAATCCCATTTTTCGACAGACATCGTCGTGCCACCTCT 1800

Qy       1801 AACACCAAAATCATAGACAGGAGCTTGTAGCTTAGCAACTATTTTATCGTC 1851
              |||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1801 AACACCAAAATCATAGACAGGAGCTTGTAGCTTAGCAACTATTTTATCGTC 1851

RESULT 2
BD063209
LOCUS       BD063209               1851 bp    DNA     linear   PAT 27-AUG-2002
DEFINITION  Mutants of streptococcal toxin A and methods of use.
ACCESSION   BD063209
VERSION     BD063209.1  GI:22608812
KEYWORDS    JP 2001505439-A/12.
SOURCE      unidentified
  ORGANISM  unidentified
            unclassified.
REFERENCE   1  (bases 1 to 1851)
  AUTHORS   Schlievert,P.M., Roggiani,M., Stoehr,J. and Ohlendorf,D.
```

```
  TITLE     Mutants of streptococcal toxin A and methods of use
  JOURNAL   Patent: JP 2001505439-A 12 24-APR-2001;
            REGENTS OF THE UNIVERSITY OF MINNESOTA
COMMENT     PN   JP 2001505439-A/12
            PD   24-APR-2001
            PF   05-DEC-1997 JP 1998525794
            PR   06-DEC-1996 US    60/032930
            PI   PATRICK M SCHLIEVERT,MANUELA ROGGIANI,JENNIFER STOEHR,DOUGLAS
            PI   OHLENDORF
            PC   C12N15/31,C07K14/315,A61K39/09,C12N1/21,C12N5/10//(C12N15/31,
            PC   C12R1:46)
            CC   Strandedness: Single;
            CC   Topology: Linear;
            FH   Key             Location/Qualifiers
            FT   CDS             828. .1583.
FEATURES            Location/Qualifiers
     source         1. .1851
                    /organism="unidentified"
                    /mol_type="genomic DNA"
                    /db_xref="taxon:32644"
ORIGIN

  Query Match           100.0%;  Score 1851;  DB 6;  Length 1851;
  Best Local Similarity 100.0%;  Pred. No. 3.2e-256;
  Matches 1851;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 CCATCACGCATCACTCATGTTTGACAGCTTATCATCGATAAGCTTACTTTTCGAATCAGG 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 CCATCACGCATCACTCATGTTTGACAGCTTATCATCGATAAGCTTACTTTTCGAATCAGG 60

Qy         61 TCTATCCTTGAAACAGGTGCAACATAGATTAGGGCATGGAGATTTACCAGACAACTATGA 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 TCTATCCTTGAAACAGGTGCAACATAGATTAGGGCATGGAGATTTACCAGACAACTATGA 120

Qy        121 ACGTATATACTCACATCACGCAATCGGCAATTGATGACATTGGAACTAAATTCAATCAAT 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 ACGTATATACTCACATCACGCAATCGGCAATTGATGACATTGGAACTAAATTCAATCAAT 180

Qy        181 TTGTTACTAACAAGCAACTAGATTGACAACTAATTCTCAACAAACGTTAATTTAACAACA 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 TTGTTACTAACAAGCAACTAGATTGACAACTAATTCTCAACAAACGTTAATTTAACAACA 240

Qy        241 TTCAAGTAACTCCCACCAGCTCCATCAATGCTTACCGTAAGTAATCATAACTTACTAAAA 300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 TTCAAGTAACTCCCACCAGCTCCATCAATGCTTACCGTAAGTAATCATAACTTACTAAAA 300

Qy        301 CCTTGTTACATCAAGGTTTTTTCTTTTTGTCTTGTTCATGAGTTACCATAACTTTCTATA 360
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        301 CCTTGTTACATCAAGGTTTTTTCTTTTTGTCTTGTTCATGAGTTACCATAACTTTCTATA 360

Qy        361 TTATTGACAACTAAATTGACAACTCTTCAATTATTTTTCTGTCTACTCAAAGTTTTCTTC 420
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        361 TTATTGACAACTAAATTGACAACTCTTCAATTATTTTTCTGTCTACTCAAAGTTTTCTTC 420

Qy        421 ATTTGATATAGTCTAATTCCACCATCACTTCTTCCACTCTCTCTACCGTCACAACTTCAT 480
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        421 ATTTGATATAGTCTAATTCCACCATCACTTCTTCCACTCTCTCTACCGTCACAACTTCAT 480

Qy        481 CATCTCTCACTTTTTCGTGTGGTAACACATAATCAAATATCTTTCCGTTTTTACGCACTA 540
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        481 CATCTCTCACTTTTTCGTGTGGTAACACATAATCAAATATCTTTCCGTTTTTACGCACTA 540

Qy        541 TCGCTACTGTGTCACCTAAAATATACCCCTTATCAATCGCTTCTTTAAACTCATCTATAT 600
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        541 TCGCTACTGTGTCACCTAAAATATACCCCTTATCAATCGCTTCTTTAAACTCATCTATAT 600

Qy        601 ATAACATATTTCATCCTCCTACCTATCTATTCGTAAAAAGATAAAAATAACTATTGTTTT 660
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        601 ATAACATATTTCATCCTCCTACCTATCTATTCGTAAAAAGATAAAAATAACTATTGTTTT 660

Qy        661 TTTTGTTATTTTATAATAAAATTATTAATATAAGTTAATGTTTTTTAAAAATATACAATT 720
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        661 TTTTGTTATTTTATAATAAAATTATTAATATAAGTTAATGTTTTTTAAAAATATACAATT 720

Qy        721 TTATTCTATTTATAGTTAGCTATTTTTTCATTGTTAGTAATATTGGTGAATTGTAATAAC 780
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        721 TTATTCTATTTATAGTTAGCTATTTTTTCATTGTTAGTAATATTGGTGAATTGTAATAAC 780

Qy        781 CTTTTTAAATCTAGAGGAGAACCCAGATATAAAATGGAGGAATATTAATGGAAAACAATA 840
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        781 CTTTTTAAATCTAGAGGAGAACCCAGATATAAAATGGAGGAATATTAATGGAAAACAATA 840

Qy        841 AAAAAGTATTGAAGAAAATGGTATTTTTTGTTTTAGTGACATTTCTTGGACTAACAATCT 900
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        841 AAAAAGTATTGAAGAAAATGGTATTTTTTGTTTTAGTGACATTTCTTGGACTAACAATCT 900

Qy        901 CGCAAGAGGTATTTGCTCAACAAGACCCCGATCCAAGCCAACTTCACAGATCTAGTTTAG 960
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        901 CGCAAGAGGTATTTGCTCAACAAGACCCCGATCCAAGCCAACTTCACAGATCTAGTTTAG 960

Qy        961 TTAAAAACCTTCAAAATATATATTTTCTTTATGAGGGTGACCCTGTTACTCACGAGAATG 1020
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        961 TTAAAAACCTTCAAAATATATATTTTCTTTATGAGGGTGACCCTGTTACTCACGAGAATG 1020

Qy       1021 TGAAATCTGTTGATCAACTTTTATCTCACCATTTAATATATAATGTTTCAGGGCCAAATT 1080
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1021 TGAAATCTGTTGATCAACTTTTATCTCACCATTTAATATATAATGTTTCAGGGCCAAATT 1080

Qy       1081 ATGATAAATTAAAAACTGAACTTAAGAACCAAGAGATGGCAACTTTATTTAAGGATAAAA 1140
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1081 ATGATAAATTAAAAACTGAACTTAAGAACCAAGAGATGGCAACTTTATTTAAGGATAAAA 1140

Qy       1141 ACGTTGATATTTATGGTGTAGAATATTACCATCTCTGTTATTTATGTGAAAATGCAGAAA 1200
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1141 ACGTTGATATTTATGGTGTAGAATATTACCATCTCTGTTATTTATGTGAAAATGCAGAAA 1200

Qy       1201 GGAGTGCATGTATCTACGGAGGGGTAACAAATCATGAAGGGAATCATTTAGAAATTCCTA 1260
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1201 GGAGTGCATGTATCTACGGAGGGGTAACAAATCATGAAGGGAATCATTTAGAAATTCCTA 1260

Qy       1261 AAAAGATAGTCGTTAAAGTATCAATCGATGGTATCCAAAGCCTATCATTTGATATTGAAA 1320
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1261 AAAAGATAGTCGTTAAAGTATCAATCGATGGTATCCAAAGCCTATCATTTGATATTGAAA 1320

Qy       1321 CAAATAAAAAAAATGGTAACTGCTCAAGAATTAGACTATAAAGTTAGAAAATATCTTACAG 1380
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1321 CAAATAAAAAAATGGTAACTGCTCAAGAATTAGACTATAAAGTTAGAAAATATCTTACAG 1380

Qy       1381 ATAATAAGCAACTATATACTAATGGACCTTCTAAATATGAAACTGGATATATAAAGTTCA 1440
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1381 ATAATAAGCAACTATATACTAATGGACCTTCTAAATATGAAACTGGATATATAAAGTTCA 1440

Qy       1441 TACCTAAGAATAAAGAAAGTTTTTGGTTTGATTTTTTCCCTGAACCAGAATTTACTCAAT 1500
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1441 TACCTAAGAATAAAGAAAGTTTTTGGTTTGATTTTTTCCCTGAACCAGAATTTACTCAAT 1500

Qy       1501 CTAAATATCTTATGATATATAAAGATAATGAAACGCTTGACTCAAACACAAGCCAAATTG 1560
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1501 CTAAATATCTTATGATATATAAAGATAATGAAACGCTTGACTCAAACACAAGCCAAATTG 1560

Qy       1561 AAGTCTACCTAACAACCAAGTAACTTTTTGCTTTTGGCAACCTTACCTACTGCTGGATTT 1620
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1561 AAGTCTACCTAACAACCAAGTAACTTTTTGCTTTTGGCAACCTTACCTACTGCTGGATTT 1620

Qy       1621 AGAAATTTTATTGCAATTCTTTTATTAATGTAAAAACCGCTCATTTGATGAGCGGTTTTG 1680
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1621 AGAAATTTTATTGCAATTCTTTTATTAATGTAAAAACCGCTCATTTGATGAGCGGTTTTG 1680

Qy       1681 TCTTATCTAAAGGAGCTTTACCTCCTAATGCTGCAAAATTTTAAATGTTGGATTTTTGTA 1740
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1681 TCTTATCTAAAGGAGCTTTACCTCCTAATGCTGCAAAATTTTAAATGTTGGATTTTTGTA 1740

Qy       1741 TTTGTCTATTGTATTTGATGGGTAATCCCATTTTTCGACAGACATCGTCGTGCCACCTCT 1800
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1741 TTTGTCTATTGTATTTGATGGGTAATCCCATTTTTCGACAGACATCGTCGTGCCACCTCT 1800
```

```
Qy       1801 AACACCAAAATCATAGACAGGAGCTTGTAGCTTAGCAACTATTTTATCGTC 1851
             ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1801 AACACCAAAATCATAGACAGGAGCTTGTAGCTTAGCAACTATTTTATCGTC 1851


RESULT 3
AR089274
LOCUS       AR089274                1837 bp    DNA     linear   PAT 07-SEP-2000
DEFINITION  Sequence 33 from patent US 5994066.
ACCESSION   AR089274
VERSION     AR089274.1  GI:10016031
KEYWORDS    .
SOURCE      Unknown.
  ORGANISM  Unknown.
            Unclassified.
REFERENCE   1  (bases 1 to 1837)
  AUTHORS   Bergeron,M.G., Picard,F.J., Ouellette,M. and Roy,P.H.
  TITLE     Species-specific and universal DNA probes and amplification primers
            to rapidly detect and identify common bacterial pathogens and
            associated antibiotic resistance genes from clinical specimens for
            routine diagnosis in microbiology laboratories
  JOURNAL   Patent: US 5994066-A 33 30-NOV-1999;
FEATURES             Location/Qualifiers
     source          1. .1837
                     /organism="unknown"
                     /mol_type="unassigned DNA"
ORIGIN

  Query Match             99.2%;  Score 1835.4;  DB 6;  Length 1837;
  Best Local Similarity   99.9%;  Pred. No. 5.6e-254;
  Matches 1836;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy         15 TCATGTTTGACAGCTTATCATCGATAAGCTTACTTTTCGAATCAGGTCTATCCTTGAAAC 74
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 TCATGTTTGACAGCTTATCATCGATAAGCTTACTTTTCGAATCAGGTCTATCCTTGAAAC 60

Qy         75 AGGTGCAACATAGATTAGGGCATGGAGATTTACCAGACAACTATGAACGTATATACTCAC 134
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 AGGTGCAACATAGATTAGGGCATGGAGATTTACCAGACAACTATGAACGTATATACTCAC 120

Qy        135 ATCACGCAATCGGCAATTGATGACATTGGAACTAAATTCAATCAATTTGTTACTAACAAG 194
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 ATCACGCAATCGGCAATTGATGACATTGGAACTAAATTCAATCAATTTGTTACTAACAAG 180

Qy        195 CAACTAGATTGACAACTAATTCTCAACAAACGTTAATTTAACAACATTCAAGTAACTCCC 254
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 CAACTAGATTGACAACTAATTCTCAACAAACGTTAATTTAACAACATTCAAGTAACTCCC 240

Qy        255 ACCAGCTCCATCAATGCTTACCGTAAGTAATCATAACTTACTAAAACCTTGTTACATCAA 314
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 ACCAGCTCCATCAATGCTTACCGTAAGTAATCATAACTTACTAAAACCTTGTTACATCAA 300

Qy        315 GGTTTTTTCTTTTTGTCTTGTTCATGAGTTACCATAACTTTCTATATTATTGACAACTAA 374
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        301 GGTTTTTTCTTTTTGTCTTGTTCATGAGTTACCATAACTTTCTATATTATTGACAACTAA 360

Qy        375 ATTGACAACTCTTCAATTATTTTTCTGTCTACTCAAAGTTTTCTTCATTTGATATAGTCT 434
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        361 ATTGACAACTCTTCAATTATTTTTCTGTCTACTCAAAGTTTTCTTCATTTGATATAGTCT 420

Qy        435 AATTCCACCATCACTTCTTCCACTCTCTCTACCGTCACAACTTCATCATCTCTCACTTTT 494
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        421 AATTCCACCATCACTTCTTCCACTCTCTCTACCGTCACAACTTCATCATCTCTCACTTTT 480

Qy        495 TCGTGTGGTAACACATAATCAAATATCTTTCCGTTTTTACGCACTATCGCTACTGTGTCA 554
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        481 TCGTGTGGTAACACATAATCAAATATCTTTCCGTTTTTACGCACTATCGCTACTGTGTCA 540

Qy        555 CCTAAAATATACCCCTTATCAATCGCTTCTTTAAACTCATCTATATATAACATATTTCAT 614
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        541 CCTAAAATATACCCCTTATCAATCGCTTCTTTAAACTCATCTATATATAACATATTTCAT 600

Qy        615 CCTCCTACCTATCTATTCGTAAAAAGATAAAAATAACTATTGTTTTTTTTGTTATTTTAT 674
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        601 CCTCCTACCTATCTATTCGTAAAAAGATAAAAATAACTATTGTTTTTTTTGTTATTTTAT 660

Qy        675 AATAAAATTATTAATATAAGTTAATGTTTTTTAAAAATATACAATTTTATTCTATTTATA 734
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        661 AATAAAATTATTAATATAAGTTAATGTTTTTTAAAAATATACAATTTTATTCTATTTATA 720

Qy        735 GTTAGCTATTTTTTCATTGTTAGTAATATTGGTGAATTGTAATAACCTTTTTAAATCTAG 794
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        721 GTTAGCTATTTTTTCATTGTTAGTAATATTGGTGAATTGTAATAACCTTTTTAAATCTAG 780

Qy        795 AGGAGAACCCAGATATAAAATGGAGGAATATTAATGGAAAACAATAAAAAAGTATTGAAG 854
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        781 AGGAGAACCCAGATATAAAATGGAGGAATATTAATGGAAAACAATAAAAAAGTATTGAAG 840

Qy        855 AAAATGGTATTTTTTGTTTTAGTGACATTTCTTGGACTAACAATCTCGCAAGAGGTATTT 914
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        841 AAAATGGTATTTTTTGTTTTAGTGACATTTCTTGGACTAACAATCTCGCAAGAGGTATTT 900

Qy        915 GCTCAACAAGACCCCGATCCAAGCCAACTTCACAGATCTAGTTTAGTTAAAAACCTTCAA 974
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        901 GCTCAACAAGACCCCGATCCAAGCCAACTTCACAGATCTAGTTTAGTTAAAAACCTTCAA 960

Qy        975 AATATATATTTTCTTTATGAGGGTGACCCTGTTACTCACGAGAATGTGAAATCTGTTGAT 1034
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        961 AATATATATTTTCTTTATGAGGGTGACCCTGTTACTCACGAGAATGTGAAATCTGTTGAT 1020

Qy       1035 CAACTTTTATCTCACCATTTAATATATAATGTTTCAGGGCCAAATTATGATAAATTAAAA 1094
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1021 CAACTTTTATCTCACGATTTAATATATAATGTTTCAGGGCCAAATTATGATAAATTAAAA 1080

Qy       1095 ACTGAACTTAAGAACCAAGAGATGGCAACTTTATTTAAGGATAAAAACGTTGATATTTAT 1154
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1081 ACTGAACTTAAGAACCAAGAGATGGCAACTTTATTTAAGGATAAAAACGTTGATATTTAT 1140

Qy       1155 GGTGTAGAATATTACCATCTCTGTTATTTATGTGAAAATGCAGAAAGGAGTGCATGTATC 1214
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1141 GGTGTAGAATATTACCATCTCTGTTATTTATGTGAAAATGCAGAAAGGAGTGCATGTATC 1200

Qy       1215 TACGGAGGGGTAACAAATCATGAAGGGAATCATTTAGAAATTCCTAAAAAGATAGTCGTT 1274
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1201 TACGGAGGGGTAACAAATCATGAAGGGAATCATTTAGAAATTCCTAAAAAGATAGTCGTT 1260

Qy       1275 AAAGTATCAATCGATGGTATCCAAAGCCTATCATTTGATATTGAAACAAATAAAAAAATG 1334
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1261 AAAGTATCAATCGATGGTATCCAAAGCCTATCATTTGATATTGAAACAAATAAAAAAATG 1320

Qy       1335 GTAACTGCTCAAGAATTAGACTATAAAGTTAGAAAATATCTTACAGATAATAAGCAACTA 1394
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1321 GTAACTGCTCAAGAATTAGACTATAAAGTTAGAAAATATCTTACAGATAATAAGCAACTA 1380

Qy       1395 TATACTAATGGACCTTCTAAATATGAAACTGGATATATAAAGTTCATACCTAAGAATAAA 1454
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1381 TATACTAATGGACCTTCTAAATATGAAACTGGATATATAAAGTTCATACCTAAGAATAAA 1440

Qy       1455 GAAAGTTTTTGGTTTGATTTTTTCCCTGAACCAGAATTTACTCAATCTAAATATCTTATG 1514
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1441 GAAAGTTTTTGGTTTGATTTTTTCCCTGAACCAGAATTTACTCAATCTAAATATCTTATG 1500

Qy       1515 ATATATAAAGATAATGAAACGCTTGACTCAAACACAAGCCAAATTGAAGTCTACCTAACA 1574
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1501 ATATATAAAGATAATGAAACGCTTGACTCAAACACAAGCCAAATTGAAGTCTACCTAACA 1560

Qy       1575 ACCAAGTAACTTTTTGCTTTTGGCAACCTTACCTACTGCTGGATTTAGAAATTTTATTGC 1634
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1561 ACCAAGTAACTTTTTGCTTTTGGCAACCTTACCTACTGCTGGATTTAGAAATTTTATTGC 1620

Qy       1635 AATTCTTTTATTAATGTAAAAACCGCTCATTTGATGAGCGGTTTTGTCTTATCTAAAGGA 1694
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1621 AATTCTTTTATTAATGTAAAAACCGCTCATTTGATGAGCGGTTTTGTCTTATCTAAAGGA 1680

Qy       1695 GCTTTACCTCCTAATGCTGCAAAATTTTAAATGTTGGATTTTTGTATTTGTCTATTGTAT 1754
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1681 GCTTTACCTCCTAATGCTGCAAAATTTTAAATGTTGGATTTTTGTATTTGTCTATTGTAT 1740
```

```
Db       1681 GCTTTACCTCCTAATGCTGCAAAATTTTAAATGTTGGATTTTTGTATTTGTCTATTGTAT 1740

Qy       1755 TTGATGGGTAATCCCATTTTTCGACAGACATCGTCGTGCCACCTCTAACACCAAAATCAT 1814
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1741 TTGATGGGTAATCCCATTTTTCGACAGACATCGTCGTGCCACCTCTAACACCAAAATCAT 1800

Qy       1815 AGACAGGAGCTTGTAGCTTAGCAACTATTTTATCGTC 1851
             |||||||||||||||||||||||||||||||||||||
Db       1801 AGACAGGAGCTTGTAGCTTAGCAACTATTTTATCGTC 1837


RESULT 4
AR093474
LOCUS       AR093474                1837 bp    DNA     linear   PAT 08-SEP-2000
DEFINITION  Sequence 33 from patent US 6001564.
ACCESSION   AR093474
VERSION     AR093474.1  GI:10020223
KEYWORDS    .
SOURCE      Unknown.
  ORGANISM  Unknown.
            Unclassified.
REFERENCE   1  (bases 1 to 1837)
  AUTHORS   Bergeron,M.G., Ouellette,M. and Roy,P.H.
  TITLE     Species specific and universal DNA probes and amplification primers
            to rapidly detect and identify common bacterial pathogens and
            associated antibiotic resistance genes from clinical specimens for
            routine diagnosis in microbiology laboratories
  JOURNAL   Patent: US 6001564-A 33 14-DEC-1999;
FEATURES             Location/Qualifiers
     source          1. .1837
                     /organism="unknown"
                     /mol_type="unassigned DNA"
ORIGIN

  Query Match           99.2%;  Score 1835.4;  DB 6;  Length 1837;
  Best Local Similarity 99.9%;  Pred. No. 5.6e-254;
  Matches 1836;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy          15 TCATGTTTGACAGCTTATCATCGATAAGCTTACTTTTCGAATCAGGTCTATCCTTGAAAC 74
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db           1 TCATGTTTGACAGCTTATCATCGATAAGCTTACTTTTCGAATCAGGTCTATCCTTGAAAC 60

Qy          75 AGGTGCAACATAGATTAGGGCATGGAGATTTACCAGACAACTATGAACGTATATACTCAC 134
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          61 AGGTGCAACATAGATTAGGGCATGGAGATTTACCAGACAACTATGAACGTATATACTCAC 120

Qy         135 ATCACGCAATCGGCAATTGATGACATTGGAACTAAATTCAATCAATTTGTTACTAACAAG 194
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         121 ATCACGCAATCGGCAATTGATGACATTGGAACTAAATTCAATCAATTTGTTACTAACAAG 180

Qy         195 CAACTAGATTGACAACTAATTCTCAACAAACGTTAATTTAACAACATTCAAGTAACTCCC 254
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         181 CAACTAGATTGACAACTAATTCTCAACAAACGTTAATTTAACAACATTCAAGTAACTCCC 240

Qy         255 ACCAGCTCCATCAATGCTTACCGTAAGTAATCATAACTTACTAAAACCTTGTTACATCAA 314
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         241 ACCAGCTCCATCAATGCTTACCGTAAGTAATCATAACTTACTAAAACCTTGTTACATCAA 300

Qy         315 GGTTTTTTCTTTTTGTCTTGTTCATGAGTTACCATAACTTTCTATATTATTGACAACTAA 374
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         301 GGTTTTTTCTTTTTGTCTTGTTCATGAGTTACCATAACTTTCTATATTATTGACAACTAA 360

Qy         375 ATTGACAACTCTTCAATTATTTTTCTGTCTACTCAAAGTTTTCTTCATTTGATATAGTCT 434
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         361 ATTGACAACTCTTCAATTATTTTTCTGTCTACTCAAAGTTTTCTTCATTTGATATAGTCT 420

Qy         435 AATTCCACCATCACTTCTTCCACTCTCTCTACCGTCACAACTTCATCATCTCTCACTTTT 494
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         421 AATTCCACCATCACTTCTTCCACTCTCTCTACCGTCACAACTTCATCATCTCTCACTTTT 480

Qy         495 TCGTGTGGTAACACATAATCAAATATCTTTCCGTTTTTACGCACTATCGCTACTGTGTCA 554
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         481 TCGTGTGGTAACACATAATCAAATATCTTTCCGTTTTTACGCACTATCGCTACTGTGTCA 540

Qy         555 CCTAAAATATACCCCTTATCAATCGCTTCTTTAAACTCATCTATATATAACATATTTCAT 614
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         541 CCTAAAATATACCCCTTATCAATCGCTTCTTTAAACTCATCTATATATAACATATTTCAT 600

Qy         615 CCTCCTACCTATCTATTCGTAAAAAGATAAAAATAACTATTGTTTTTTTTGTTATTTTAT 674
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         601 CCTCCTACCTATCTATTCGTAAAAAGATAAAAATAACTATTGTTTTTTTTGTTATTTTAT 660

Qy         675 AATAAAATTATTAATATAAGTTAATGTTTTTTAAAAATATACAATTTTATTCTATTTATA 734
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         661 AATAAAATTATTAATATAAGTTAATGTTTTTTAAAAATATACAATTTTATTCTATTTATA 720

Qy         735 GTTAGCTATTTTTTCATTGTTAGTAATATTGGTGAATTGTAATAACCTTTTTAAATCTAG 794
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         721 GTTAGCTATTTTTTCATTGTTAGTAATATTGGTGAATTGTAATAACCTTTTTAAATCTAG 780

Qy         795 AGGAGAACCCAGATATAAAATGGAGGAATATTAATGGAAAACAATAAAAAAAGTATTGAAG 854
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         781 AGGAGAACCCAGATATAAAATGGAGGAATATTAATGGAAAACAATAAAAAAAGTATTGAAG 840

Qy         855 AAAATGGTATTTTTTGTTTTAGTGACATTTCTTGGACTAACAATCTCGCAAGAGGTATTT 914
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         841 AAAATGGTATTTTTTGTTTTAGTGACATTTCTTGGACTAACAATCTCGCAAGAGGTATTT 900

Qy         915 GCTCAACAAGACCCCGATCCAAGCCAACTTCACAGATCTAGTTTAGTTAAAAACCTTCAA 974
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         901 GCTCAACAAGACCCCGATCCAAGCCAACTTCACAGATCTAGTTTAGTTAAAAACCTTCAA 960

Qy         975 AATATATATTTTCTTTATGAGGGTGACCCTGTTACTCACGAGAATGTGAAATCTGTTGAT 1034
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         961 AATATATATTTTCTTTATGAGGGTGACCCTGTTACTCACGAGAATGTGAAATCTGTTGAT 1020

Qy        1035 CAACTTTTATCTCACCATTTAATATATAATGTTTCAGGGCCAAATTATGATAAATTAAAA 1094
               ||||||||||||||| ||||||||||||||||||||||||||||||||||||||||||||
Db        1021 CAACTTTTATCTCACGATTTAATATATAATGTTTCAGGGCCAAATTATGATAAATTAAAA 1080

Qy        1095 ACTGAACTTAAGAACCAAGAGATGGCAACTTTATTTAAGGATAAAAACGTTGATATTTAT 1154
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        1081 ACTGAACTTAAGAACCAAGAGATGGCAACTTTATTTAAGGATAAAAACGTTGATATTTAT 1140

Qy        1155 GGTGTAGAATATTACCATCTCTGTTATTTATGTGAAAATGCAGAAAGGAGTGCATGTATC 1214
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        1141 GGTGTAGAATATTACCATCTCTGTTATTTATGTGAAAATGCAGAAAGGAGTGCATGTATC 1200

Qy        1215 TACGGAGGGGTAACAAATCATGAAGGGAATCATTTAGAAATTCCTAAAAAGATAGTCGTT 1274
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        1201 TACGGAGGGGTAACAAATCATGAAGGGAATCATTTAGAAATTCCTAAAAAGATAGTCGTT 1260

Qy        1275 AAAGTATCAATCGATGGTATCCAAAGCCTATCATTTGATATTGAAACAAATAAAAAAATG 1334
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        1261 AAAGTATCAATCGATGGTATCCAAAGCCTATCATTTGATATTGAAACAAATAAAAAAATG 1320

Qy        1335 GTAACTGCTCAAGAATTAGACTATAAAGTTAGAAAATATCTTACAGATAATAAGCAACTA 1394
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        1321 GTAACTGCTCAAGAATTAGACTATAAAGTTAGAAAATATCTTACAGATAATAAGCAACTA 1380

Qy        1395 TATACTAATGGACCTTCTAAATATGAAACTGGATATATAAAGTTCATACCTAAGAATAAA 1454
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        1381 TATACTAATGGACCTTCTAAATATGAAACTGGATATATAAAGTTCATACCTAAGAATAAA 1440

Qy        1455 GAAAGTTTTTGGTTTGATTTTTTCCCTGAACCAGAATTTACTCAATCTAAATATCTTATG 1514
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        1441 GAAAGTTTTTGGTTTGATTTTTTCCCTGAACCAGAATTTACTCAATCTAAATATCTTATG 1500

Qy        1515 ATATATAAAGATAATGAAACGCTTGACTCAAACACAAGCCAAATTGAAGTCTACCTAACA 1574
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        1501 ATATATAAAGATAATGAAACGCTTGACTCAAACACAAGCCAAATTGAAGTCTACCTAACA 1560

Qy        1575 ACCAAGTAACTTTTTGCTTTTGGCAACCTTACCTACTGCTGGATTTAGAAATTTTATTGC 1634
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        1561 ACCAAGTAACTTTTTGCTTTTGGCAACCTTACCTACTGCTGGATTTAGAAATTTTATTGC 1620
```

```
Qy       1635 AATTCTTTTATTAATGTAAAAACCGCTCATTTGATGAGCGGTTTTGTCTTATCTAAAGGA 1694
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1621 AATTCTTTTATTAATGTAAAAACCGCTCATTTGATGAGCGGTTTTGTCTTATCTAAAGGA 1680

Qy       1695 GCTTTACCTCCTAATGCTGCAAAATTTTAAATGTTGGATTTTTGTATTTGTCTATTGTAT 1754
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1681 GCTTTACCTCCTAATGCTGCAAAATTTTAAATGTTGGATTTTTGTATTTGTCTATTGTAT 1740

Qy       1755 TTGATGGGTAATCCCATTTTTCGACAGACATCGTCGTGCCACCTCTAACACCAAAATCAT 1814
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1741 TTGATGGGTAATCCCATTTTTCGACAGACATCGTCGTGCCACCTCTAACACCAAAATCAT 1800

Qy       1815 AGACAGGAGCTTGTAGCTTAGCAACTATTTTATCGTC 1851
             |||||||||||||||||||||||||||||||||||||
Db       1801 AGACAGGAGCTTGTAGCTTAGCAACTATTTTATCGTC 1837
```

RESULT 5
AR211577
LOCUS AR211577 1837 bp DNA linear PAT 20-JUN-2002
DEFINITION Sequence 15 from patent US 6399332.
ACCESSION AR211577
VERSION AR211577.1 GI:21514936
KEYWORDS .
SOURCE Unknown.
ORGANISM Unknown.
Unclassified.
REFERENCE 1 (bases 1 to 1837)
AUTHORS Ulrich,R.G., Olson,M.A. and Bavari,S.
TITLE Bacterial superantigen vaccines
JOURNAL Patent: US 6399332-A 15 04-JUN-2002;
FEATURES Location/Qualifiers
source 1. .1837
/organism="unknown"
/mol_type="unassigned DNA"
ORIGIN

Query Match 99.0%; Score 1832.2; DB 6; Length 1837;
Best Local Similarity 99.8%; Pred. No. 1.6e-253;
Matches 1834; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

```
Qy         15 TCATGTTTGACAGCTTATCATCGATAAGCTTACTTTTCGAATCAGGTCTATCCTTGAAAC 74
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 TCATGTTTGACAGCTTATCATCGATAAGCTTACTTTTCGAATCAGGTCTATCCTTGAAAC 60

Qy         75 AGGTGCAACATAGATTAGGGCATGGAGATTTACCAGACAACTATGAACGTATATACTCAC 134
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 AGGTGCAACATAGATTAGGGCATGGAGATTTACCAGACAACTATGAACGTATATACTCAC 120

Qy        135 ATCACGCAATCGGCAATTGATGACATTGGAACTAAATTCAATCAATTTGTTACTAACAAG 194
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 ATCACGCAATCGGCAATTGATGACATTGGAACTAAATTCAATCAATTTGTTACTAACAAG 180

Qy        195 CAACTAGATTGACAACTAATTCTCAACAAACGTTAATTTAACAACATTCAAGTAACTCCC 254
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 CAACTAGATTGACAACTAATTCTCAACAAACGTTAATTTAACAACATTCAAGTAACTCCC 240

Qy        255 ACCAGCTCCATCAATGCTTACCGTAAGTAATCATAACTTACTAAAACCTTGTTACATCAA 314
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 ACCAGCTCCATCAATGCTTACCGTAAGTAATCATAACTTACTAAAACCTTGTTACATCAA 300

Qy        315 GGTTTTTTCTTTTTGTCTTGTTCATGAGTTACCATAACTTTCTATATTATTGACAACTAA 374
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        301 GGTTTTTTCTTTTTGTCTTGTTCATGAGTTACCATAACTTTCTATATTATTGACAACTAA 360

Qy        375 ATTGACAACTCTTCAATTATTTTTCTGTCTACTCAAAGTTTTCTTCATTTGATATAGTCT 434
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        361 ATTGACAACTCTTCAATTATTTTTCTGTCTACTCAAAGTTTTCTTCATTTGATATAGTCT 420

Qy        435 AATTCCACCATCACTTCTTCCACTCTCTCTACCGTCACAACTTCATCATCTCTCACTTTT 494
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        421 AATTCCACCATCACTTCTTCCACTCTCTCTACCGTCACAACTTCATCATCTCTCACTTTT 480
```

```
Qy        495 TCGTGTGGTAACACATAATCAAATATCTTTCCGTTTTTACGCACTATCGCTACTGTGTCA 554
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        481 TCGTGTGGTAACACATAATCAAATATCTTTCCGTTTTTACGCACTATCGCTACTGTGTCA 540

Qy        555 CCTAAAATATACCCCTTATCAATCGCTTCTTTAAACTCATCTATATATAACATATTTCAT 614
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        541 CCTAAAATATACCCCTTATCAATCGCTTCTTTAAACTCATCTATATATAACATATTTCAT 600

Qy        615 CCTCCTACCTATCTATTCGTAAAAAGATAAAAATAACTATTGTTTTTTTTTGTTATTTTAT 674
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        601 CCTCCTACCTATCTATTCGTAAAAAGATAAAAATAACTATTGTTTTTTTTTGTTATTTTAT 660

Qy        675 AATAAAATTATTAATATAAGTTAATGTTTTTTAAAAATATACAATTTTATTCTATTTATA 734
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        661 AATAAAATTATTAATATAAGTTAATGTTTTTTAAAAATATACAATTTTATTCTATTTATA 720

Qy        735 GTTAGCTATTTTTTCATTGTTAGTAATATTGGTGAATTGTAATAACCTTTTTAAATCTAG 794
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        721 GTTAGCTATTTTTTCATTGTTAGTAATATTGGTGAATTGTAATAACCTTTTTAAATCTAG 780

Qy        795 AGGAGAACCCAGATATAAAATGGAGGAATATTAATGGAAAACAATAAAAAAGTATTGAAG 854
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        781 AGGAGAACCCAGATATAAAATGGAGGAATATTAATGGAAAACAATAAAAAAGTATTGAAG 840

Qy        855 AAAATGGTATTTTTTGTTTTAGTGACATTTCTTGGACTAACAATCTCGCAAGAGGTATTT 914
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        841 AAAATGGTATTTTTTGTTTTAGTGACATTTCTTGGACTAACAATCTCGCAAGAGGTATTT 900

Qy        915 GCTCAACAAGACCCCGATCCAAGCCAACTTCACAGATCTAGTTTAGTTAAAAACCTTCAA 974
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        901 GCTCAACAAGACCCCGATCCAAGCCAACTTCACAGATCTAGTTTAGTTAAAAACCTTCAA 960

Qy        975 AATATATATTTTCTTTATGAGGGTGACCCTGTTACTCACGAGAATGTGAAATCTGTTGAT 1034
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        961 AATATATATTTTCTTTATGAGGGTGACCCTGTTACTCACGAGAATGTGAAATCTGTTGAT 1020

Qy       1035 CAACTTTTATCTCACCATTTAATATATAATGTTTCAGGGCCAAATTATGATAAATTAAAA 1094
              |||||| |||||| ||||||||||||||||||||||||||||||||||||||||||||||
Db       1021 CAACTTAGATCTCACGATTTAATATATAATGTTTCAGGGCCAAATTATGATAAATTAAAA 1080

Qy       1095 ACTGAACTTAAGAACCAAGAGATGGCAACTTTATTTAAGGATAAAAACGTTGATATTTAT 1154
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1081 ACTGAACTTAAGAACCAAGAGATGGCAACTTTATTTAAGGATAAAAACGTTGATATTTAT 1140

Qy       1155 GGTGTAGAATATTACCATCTCTGTTATTTATGTGAAAATGCAGAAAGGAGTGCATGTATC 1214
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1141 GGTGTAGAATATTACCATCTCTGTTATTTATGTGAAAATGCAGAAAGGAGTGCATGTATC 1200

Qy       1215 TACGGAGGGGTAACAAATCATGAAGGGAATCATTTAGAAATTCCTAAAAAGATAGTCGTT 1274
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1201 TACGGAGGGGTAACAAATCATGAAGGGAATCATTTAGAAATTCCTAAAAAGATAGTCGTT 1260

Qy       1275 AAAGTATCAATCGATGGTATCCAAAGCCTATCATTTGATATTGAAACAAATAAAAAAATG 1334
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1261 AAAGTATCAATCGATGGTATCCAAAGCCTATCATTTGATATTGAAACAAATAAAAAAATG 1320

Qy       1335 GTAACTGCTCAAGAATTAGACTATAAAGTTAGAAAATATCTTACAGATAATAAGCAACTA 1394
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1321 GTAACTGCTCAAGAATTAGACTATAAAGTTAGAAAATATCTTACAGATAATAAGCAACTA 1380

Qy       1395 TATACTAATGGACCTTCTAAATATGAAACTGGATATATAAAGTTCATACCTAAGAATAAA 1454
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1381 TATACTAATGGACCTTCTAAATATGAAACTGGATATATAAAGTTCATACCTAAGAATAAA 1440

Qy       1455 GAAAGTTTTTGGTTTGATTTTTTCCCTGAACCAGAATTTACTCAATCTAAATATCTTATG 1514
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1441 GAAAGTTTTTGGTTTGATTTTTTCCCTGAACCAGAATTTACTCAATCTAAATATCTTATG 1500

Qy       1515 ATATATAAAGATAATGAAACGCTTGACTCAAACACAAGCCAAATTGAAGTCTACCTAACA 1574
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1501 ATATATAAAGATAATGAAACGCTTGACTCAAACACAAGCCAAATTGAAGTCTACCTAACA 1560

Qy       1575 ACCAAGTAACTTTTTGCTTTTGGCAACCTTACCTACTGCTGGATTTAGAAATTTTATTGC 1634
```

```
Db      1561 ACCAAGTAACTTTTTGCTTTTGGCAACCTTACCTACTGCTGGATTTAGAAATTTTATTGC 1620

Qy      1635 AATTCTTTTATTAATGTAAAAACCGCTCATTTGATGAGCGGTTTTGTCTTATCTAAAGGA 1694
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      1621 AATTCTTTTATTAATGTAAAAACCGCTCATTTGATGAGCGGTTTTGTCTTATCTAAAGGA 1680

Qy      1695 GCTTTACCTCCTAATGCTGCAAAATTTTAAATGTTGGATTTTTGTATTTGTCTATTGTAT 1754
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      1681 GCTTTACCTCCTAATGCTGCAAAATTTTAAATGTTGGATTTTTGTATTTGTCTATTGTAT 1740

Qy      1755 TTGATGGGTAATCCCATTTTTCGACAGACATCGTCGTGCCACCTCTAACACCAAAATCAT 1814
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      1741 TTGATGGGTAATCCCATTTTTCGACAGACATCGTCGTGCCACCTCTAACACCAAAATCAT 1800

Qy      1815 AGACAGGAGCTTGTAGCTTAGCAACTATTTTATCGTC 1851
             |||||||||||||||||||||||||||||||||||||
Db      1801 AGACAGGAGCTTGTAGCTTAGCAACTATTTTATCGTC 1837


RESULT 6
BD223209
LOCUS       BD223209               1837 bp    DNA     linear   PAT 17-JUL-2003
DEFINITION  Bacterial superantigen vaccine.
ACCESSION   BD223209
VERSION     BD223209.1  GI:33032979
KEYWORDS    JP 2002522055-A/8.
SOURCE      unidentified
  ORGANISM  unidentified
            unclassified.
REFERENCE   1  (bases 1 to 1837)
  AUTHORS   Ulrich,R.G., Olson,M.A. and Bavari,S.
  TITLE     Bacterial superantigen vaccine
  JOURNAL   Patent: JP 2002522055-A 8 23-JUL-2002;
            WALTER REED ARMY INSTITUTE OF RESEARCH
COMMENT     OS   Unidentified
            PN   JP 2002522055-A/8
            PD   23-JUL-2002
            PF   13-AUG-1998 JP 2000564656
            PI   ROBERT G ULRICH,MARK A OLSON,SINA BAVARI
            PC   C12N15/09,A61K39/085,A61K39/09,A61P31/00,A61P35/00,A61P37/04,
            PC   C07K14/31,
            PC   C12N1/21,C12P21/02,G01N33/53,G01N33/569,C12N15/00 CC
            Strandedness: Unknown;
            CC   Topology: Unknown;
            CC   Bacterial superantigen vaccine
            FH   Key             Location/Qualifiers
            FT   source          1. .1837
            FT                   /organism='Unidentified'.
FEATURES             Location/Qualifiers
     source          1. .1837
                     /organism="unidentified"
                     /mol_type="genomic DNA"
                     /db_xref="taxon:32644"
ORIGIN

  Query Match             99.0%;  Score 1832.2;  DB 6;  Length 1837;
  Best Local Similarity   99.8%;  Pred. No. 1.6e-253;
  Matches 1834;  Conservative    0;  Mismatches    3;  Indels    0;  Gaps    0;

Qy          15 TCATGTTTGACAGCTTATCATCGATAAGCTTACTTTTCGAATCAGGTCTATCCTTGAAAC 74
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db           1 TCATGTTTGACAGCTTATCATCGATAAGCTTACTTTTCGAATCAGGTCTATCCTTGAAAC 60

Qy          75 AGGTGCAACATAGATTAGGGCATGGAGATTTACCAGACAACTATGAACGTATATACTCAC 134
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          61 AGGTGCAACATAGATTAGGGCATGGAGATTTACCAGACAACTATGAACGTATATACTCAC 120

Qy         135 ATCACGCAATCGGCAATTGATGACATTGGAACTAAATTCAATCAATTTGTTACTAACAAG 194
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         121 ATCACGCAATCGGCAATTGATGACATTGGAACTAAATTCAATCAATTTGTTACTAACAAG 180

Qy         195 CAACTAGATTGACAACTAATTCTCAACAAACGTTAATTTAACAACATTCAAGTAACTCCC 254
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         181 CAACTAGATTGACAACTAATTCTCAACAAACGTTAATTTAACAACATTCAAGTAACTCCC 240

Qy         255 ACCAGCTCCATCAATGCTTACCGTAAGTAATCATAACTTACTAAAACCTTGTTACATCAA 314
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         241 ACCAGCTCCATCAATGCTTACCGTAAGTAATCATAACTTACTAAAACCTTGTTACATCAA 300

Qy         315 GGTTTTTTCTTTTTGTCTTGTTCATGAGTTACCATAACTTTCTATATTATTGACAACTAA 374
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         301 GGTTTTTTCTTTTTGTCTTGTTCATGAGTTACCATAACTTTCTATATTATTGACAACTAA 360

Qy         375 ATTGACAACTCTTCAATTATTTTTCTGTCTACTCAAAGTTTTCTTCATTTGATATAGTCT 434
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         361 ATTGACAACTCTTCAATTATTTTTCTGTCTACTCAAAGTTTTCTTCATTTGATATAGTCT 420

Qy         435 AATTCCACCATCACTTCTTCCACTCTCTCTACCGTCACAACTTCATCATCTCTCACTTTT 494
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         421 AATTCCACCATCACTTCTTCCACTCTCTCTACCGTCACAACTTCATCATCTCTCACTTTT 480

Qy         495 TCGTGTGGTAACACATAATCAAATATCTTTCCGTTTTTACGCACTATCGCTACTGTGTCA 554
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         481 TCGTGTGGTAACACATAATCAAATATCTTTCCGTTTTTACGCACTATCGCTACTGTGTCA 540

Qy         555 CCTAAAATATACCCCTTATCAATCGCTTCTTTAAACTCATCTATATATAACATATTTCAT 614
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         541 CCTAAAATATACCCCTTATCAATCGCTTCTTTAAACTCATCTATATATAACATATTTCAT 600

Qy         615 CCTCCTACCTATCTATTCGTAAAAAGATAAAAATAACTATTGTTTTTTTTGTTATTTTAT 674
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         601 CCTCCTACCTATCTATTCGTAAAAAGATAAAAATAACTATTGTTTTTTTTGTTATTTTAT 660

Qy         675 AATAAAATTATTAATATAAGTTAATGTTTTTTAAAAATATACAATTTTATTCTATTTATA 734
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         661 AATAAAATTATTAATATAAGTTAATGTTTTTTAAAAATATACAATTTTATTCTATTTATA 720

Qy         735 GTTAGCTATTTTTTCATTGTTAGTAATATTGGTGAATTGTAATAACCTTTTTAAATCTAG 794
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         721 GTTAGCTATTTTTTCATTGTTAGTAATATTGGTGAATTGTAATAACCTTTTTAAATCTAG 780

Qy         795 AGGAGAACCCAGATATAAAATGGAGGAATATTAATGGAAAACAATAAAAAAGTATTGAAG 854
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         781 AGGAGAACCCAGATATAAAATGGAGGAATATTAATGGAAAACAATAAAAAAGTATTGAAG 840

Qy         855 AAAATGGTATTTTTTGTTTTAGTGACATTTCTTGGACTAACAATCTCGCAAGAGGTATTT 914
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         841 AAAATGGTATTTTTTGTTTTAGTGACATTTCTTGGACTAACAATCTCGCAAGAGGTATTT 900

Qy         915 GCTCAACAAGACCCCGATCCAAGCCAACTTCACAGATCTAGTTTAGTTAAAAACCTTCAA 974
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         901 GCTCAACAAGACCCCGATCCAAGCCAACTTCACAGATCTAGTTTAGTTAAAAACCTTCAA 960

Qy         975 AATATATATTTTCTTTATGAGGGTGACCCTGTTACTCACGAGAATGTGAAATCTGTTGAT 1034
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         961 AATATATATTTTCTTTATGAGGGTGACCCTGTTACTCACGAGAATGTGAAATCTGTTGAT 1020

Qy        1035 CAACTTTTATCTCACCATTTAATATATAATGTTTCAGGGCCAAATTATGATAAATTAAAA 1094
               |||||| ||||||| ||||||||||||||||||||||||||||||||||||||||||||
Db        1021 CAACTTAGATCTCACGATTTAATATATAATGTTTCAGGGCCAAATTATGATAAATTAAAA 1080

Qy        1095 ACTGAACTTAAGAACCAAGAGATGGCAACTTTATTTAAGGATAAAAACGTTGATATTTAT 1154
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        1081 ACTGAACTTAAGAACCAAGAGATGGCAACTTTATTTAAGGATAAAAACGTTGATATTTAT 1140

Qy        1155 GGTGTAGAATATTACCATCTCTGTTATTTATGTGAAAATGCAGAAAGGAGTGCATGTATC 1214
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        1141 GGTGTAGAATATTACCATCTCTGTTATTTATGTGAAAATGCAGAAAGGAGTGCATGTATC 1200

Qy        1215 TACGGAGGGGTAACAAATCATGAAGGGAATCATTTAGAAATTCCTAAAAAGATAGTCGTT 1274
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        1201 TACGGAGGGGTAACAAATCATGAAGGGAATCATTTAGAAATTCCTAAAAAGATAGTCGTT 1260

Qy        1275 AAAGTATCAATCGATGGTATCCAAAGCCTATCATTTGATATTGAAACAAATAAAAAAATG 1334
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
```

```
Db       1261 AAAGTATCAATCGATGGTATCCAAAGCCTATCATTTGATATTGAAACAAATAAAAAAATG 1320
Qy       1335 GTAACTGCTCAAGAATTAGACTATAAAGTTAGAAAATATCTTACAGATAATAAGCAACTA 1394
Db       1321 GTAACTGCTCAAGAATTAGACTATAAAGTTAGAAAATATCTTACAGATAATAAGCAACTA 1380
Qy       1395 TATACTAATGGACCTTCTAAATATGAAACTGGATATATAAAGTTCATACCTAAGAATAAA 1454
Db       1381 TATACTAATGGACCTTCTAAATATGAAACTGGATATATAAAGTTCATACCTAAGAATAAA 1440
Qy       1455 GAAAGTTTTTGGTTTGATTTTTTCCCTGAACCAGAATTTACTCAATCTAAATATCTTATG 1514
Db       1441 GAAAGTTTTTGGTTTGATTTTTTCCCTGAACCAGAATTTACTCAATCTAAATATCTTATG 1500
Qy       1515 ATATATAAAGATAATGAAACGCTTGACTCAAACACAAGCCAAATTGAAGTCTACCTAACA 1574
Db       1501 ATATATAAAGATAATGAAACGCTTGACTCAAACACAAGCCAAATTGAAGTCTACCTAACA 1560
Qy       1575 ACCAAGTAACTTTTTGCTTTTGGCAACCTTACCTACTGCTGGATTTAGAAATTTTATTGC 1634
Db       1561 ACCAAGTAACTTTTTGCTTTTGGCAACCTTACCTACTGCTGGATTTAGAAATTTTATTGC 1620
Qy       1635 AATTCTTTTATTAATGTAAAAACCGCTCATTTGATGAGCGGTTTTGTCTTATCTAAAGGA 1694
Db       1621 AATTCTTTTATTAATGTAAAAACCGCTCATTTGATGAGCGGTTTTGTCTTATCTAAAGGA 1680
Qy       1695 GCTTTACCTCCTAATGCTGCAAAATTTTAAATGTTGGATTTTTGTATTTGTCTATTGTAT 1754
Db       1681 GCTTTACCTCCTAATGCTGCAAAATTTTAAATGTTGGATTTTTGTATTTGTCTATTGTAT 1740
Qy       1755 TTGATGGGTAATCCCATTTTTCGACAGACATCGTCGTGCCACCTCTAACACCAAAATCAT 1814
Db       1741 TTGATGGGTAATCCCATTTTTCGACAGACATCGTCGTGCCACCTCTAACACCAAAATCAT 1800
Qy       1815 AGACAGGAGCTTGTAGCTTAGCAACTATTTTATCGTC 1851
Db       1801 AGACAGGAGCTTGTAGCTTAGCAACTATTTTATCGTC 1837
```

RESULT 7
AX800046
LOCUS AX800046 1837 bp DNA linear PAT 13-OCT-2003
DEFINITION Sequence 15 from Patent WO03056015.
ACCESSION AX800046
VERSION AX800046.1 GI:37653325
KEYWORDS .
SOURCE synthetic construct
ORGANISM synthetic construct
artificial sequences.
REFERENCE 1
AUTHORS Ulrich,R.G.
TITLE Bacterial superantigen vaccines
JOURNAL Patent: WO 03056015-A 15 10-JUL-2003;
U.S. Medical Research Institute of Infectious Diseases (US)
FEATURES Location/Qualifiers
source 1. .1837
/organism="synthetic construct"
/mol_type="unassigned DNA"
/db_xref="taxon:32630"
/note="streptococcal pyrogenic exotoxin-A mutant"
ORIGIN

Query Match 98.9%; Score 1830.6; DB 6; Length 1837;
Best Local Similarity 99.8%; Pred. No. 2.7e-253;
Matches 1833; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

```
Qy         15 TCATGTTTGACAGCTTATCATCGATAAGCTTACTTTTCGAATCAGGTCTATCCTTGAAAC 74
Db          1 TCATGTTTGACAGCTTATCATCGATAAGCTTACTTTTCGAATCAGGTCTATCCTTGAAAC 60
Qy         75 AGGTGCAACATAGATTAGGGCATGGAGATTTACCAGACAACTATGAACGTATATACTCAC 134
Db         61 AGGTGCAACATAGATTAGGGCATGGAGATTTACCAGACAACTATGAACGTATATACTCAC 120
Qy        135 ATCACGCAATCGGCAATTGATGACATTGGAACTAAATTCAATCAATTTGTTACTAACAAG 194
Db        121 ATCACGCAATCGGCAATTGATGACATTGGAACTAAATTCAATCAATTTGTTACTAACAAG 180
Qy        195 CAACTAGATTGACAACTAATTCTCAACAAACGTTAATTTAACAACATTCAAGTAACTCCC 254
Db        181 CAACTAGATTGACAACTAATTCTCAACAAACGTTAATTTAACAACATTCAAGTAACTCCC 240
Qy        255 ACCAGCTCCATCAATGCTTACCGTAAGTAATCATAACTTACTAAAACCTTGTTACATCAA 314
Db        241 ACCAGCTCCATCAATGCTTACCGTAAGTAATCATAACTTACTAAAACCTTGTTACATCAA 300
Qy        315 GGTTTTTTCTTTTTGTCTTGTTCATGAGTTACCATAACTTTCTATATTATTGACAACTAA 374
Db        301 GGTTTTTTCTTTTTGTCTTGTTCATGAGTTACCATAACTTTCTATATTATTGACAACTAA 360
Qy        375 ATTGACAACTCTTCAATTATTTTTCTGTCTACTCAAAGTTTTCTTCATTTGATATAGTCT 434
Db        361 ATTGACAACTCTTCAATTATTTTTCTGTCTACTCAAAGTTTTCTTCATTTGATATAGTCT 420
Qy        435 AATTCCACCATCACTTCTTCCACTCTCTCTACCGTCACAACTTCATCATCTCTCACTTTT 494
Db        421 AATTCCACCATCACTTCTTCCACTCTCTCTACCGTCACAACTTCATCATCTCTCACTTTT 480
Qy        495 TCGTGTGGTAACACATAATCAAATATCTTTCCGTTTTTACGCACTATCGCTACTGTGTCA 554
Db        481 TCGTGTGGTAACACATAATCAAATATCTTTCCGTTTTTACGCACTATCGCTACTGTGTCA 540
Qy        555 CCTAAAATATACCCCTTATCAATCGCTTCTTTAAACTCATCTATATATAACATATTTCAT 614
Db        541 CCTAAAATATACCCCTTATCAATCGCTTCTTTAAACTCATCTATATATAACATATTTCAT 600
Qy        615 CCTCCTACCTATCTATTCGTAAAAAGATAAAAATAACTATTGTTTTTTTTGTTATTTTAT 674
Db        601 CCTCCTACCTATCTATTCGTAAAAAGATAAAAATAACTATTGTTTTTTTTGTTATTTTAT 660
Qy        675 AATAAAATTATTAATATAAGTTAATGTTTTTTAAAAATATACAATTTTATTCTATTTATA 734
Db        661 AATAAAATTATTAATATAAGTTAATGTTTTTTAAAAATATACAATTTTATTCTATTTATA 720
Qy        735 GTTAGCTATTTTTTCATTGTTAGTAATATTGGTGAATTGTAATAACCTTTTTAAATCTAG 794
Db        721 GTTAGCTATTTTTTCATTGTTAGTAATATTGGTGAATTGTAATAACCTTTTTAAATCTAG 780
Qy        795 AGGAGAACCCAGATATAAAATGGAGGAATATTAATGGAAAACAATAAAAAAGTATTGAAG 854
Db        781 AGGAGAACCCAGATATAAAATGGAGGAATATTAATGGAAAACAATAAAAAAGTATTGAAG 840
Qy        855 AAAATGGTATTTTTTGTTTTAGTGACATTTCTTGGACTAACAATCTCGCAAGAGGTATTT 914
Db        841 AAAATGGTATTTTTTGTTTTAGTGACATTTCTTGGACTAACAATCTCGCAAGAGGTATTT 900
Qy        915 GCTCAACAAGACCCCGATCCAAGCCAACTTCACAGATCTAGTTTAGTTAAAAACCTTCAA 974
Db        901 GCTCAACAAGACCCCGATCCAAGCCAACTTCACAGATCTAGTTTAGTTAAAAACCTTCAA 960
Qy        975 AATATATATTTTCTTTATGAGGGTGACCCTGTTACTCACGAGAATGTGAAATCTGTTGAT 1034
Db        961 AATATATATTTTCTTTATGAGGGTGACCCTGTTACTCACGAGAATGTGAAATCTGTTGAT 1020
Qy       1035 CAACTTTTATCTCACCATTTAATATATAATGTTTCAGGGCCAAATTATGATAAATTAAAA 1094
Db       1021 CAACTTAGATCTCACGATTTAATATATAATGTTTCAGGGCCAAATTATGATAAATTAAAA 1080
Qy       1095 ACTGAACTTAAGAACCAAGAGATGGCAACTTTATTTAAGGATAAAAACGTTGATATTTAT 1154
Db       1081 ACTGAACTTAAGAACCAAGAGATGGCAACTTTATTTAAGGATAAAAACGTTGATATTTAT 1140
Qy       1155 GGTGTAGAATATTACCATCTCTGTTATTTATGTGAAAATGCAGAAAGGAGTGCATGTATC 1214
Db       1141 GGTGTAGAATATTACCATCTCTGTTATTTATGTGAAAATGCAGAAAGGAGTGCATGTATC 1200
```

```
Qy       1215 TACGGAGGGGTAACAAATCATGAAGGGAATCATTTAGAAATTCCTAAAAAGATAGTCGTT 1274
             ||||||||||||||||||||||||||| ||||||||||||||||||||||||||||||||||
Db       1201 TACGGAGGGGTAACAAATCATGAAGGAAATCATTTAGAAATTCCTAAAAAGATAGTCGTT 1260

Qy       1275 AAAGTATCAATCGATGGTATCCAAAGCCTATCATTTGATATTGAAACAAATAAAAAAATG 1334
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1261 AAAGTATCAATCGATGGTATCCAAAGCCTATCATTTGATATTGAAACAAATAAAAAAATG 1320

Qy       1335 GTAACTGCTCAAGAATTAGACTATAAAGTTAGAAAATATCTTACAGATAATAAGCAACTA 1394
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1321 GTAACTGCTCAAGAATTAGACTATAAAGTTAGAAAATATCTTACAGATAATAAGCAACTA 1380

Qy       1395 TATACTAATGGACCTTCTAAATATGAAACTGGATATATAAAGTTCATACCTAAGAATAAA 1454
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1381 TATACTAATGGACCTTCTAAATATGAAACTGGATATATAAAGTTCATACCTAAGAATAAA 1440

Qy       1455 GAAAGTTTTTGGTTTGATTTTTTCCCTGAACCAGAATTTACTCAATCTAAATATCTTATG 1514
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1441 GAAAGTTTTTGGTTTGATTTTTTCCCTGAACCAGAATTTACTCAATCTAAATATCTTATG 1500

Qy       1515 ATATATAAAGATAATGAAACGCTTGACTCAAACACAAGCCAAATTGAAGTCTACCTAACA 1574
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1501 ATATATAAAGATAATGAAACGCTTGACTCAAACACAAGCCAAATTGAAGTCTACCTAACA 1560

Qy       1575 ACCAAGTAACTTTTTGCTTTTGGCAACCTTACCTACTGCTGGATTTAGAAATTTTATTGC 1634
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1561 ACCAAGTAACTTTTTGCTTTTGGCAACCTTACCTACTGCTGGATTTAGAAATTTTATTGC 1620

Qy       1635 AATTCTTTTATTAATGTAAAAACCGCTCATTTGATGAGCGGTTTTGTCTTATCTAAAGGA 1694
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1621 AATTCTTTTATTAATGTAAAAACCGCTCATTTGATGAGCGGTTTTGTCTTATCTAAAGGA 1680

Qy       1695 GCTTTACCTCCTAATGCTGCAAAATTTTAAATGTTGGATTTTTGTATTTGTCTATTGTAT 1754
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1681 GCTTTACCTCCTAATGCTGCAAAATTTTAAATGTTGGATTTTTGTATTTGTCTATTGTAT 1740

Qy       1755 TTGATGGGTAATCCCATTTTTCGACAGACATCGTCGTGCCACCTCTAACACCAAAATCAT 1814
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1741 TTGATGGGTAATCCCATTTTTCGACAGACATCGTCGTGCCACCTCTAACACCAAAATCAT 1800

Qy       1815 AGACAGGAGCTTGTAGCTTAGCAACTATTTTATCGTC 1851
             |||||||||||||||||||||||||||||||||||||
Db       1801 AGACAGGAGCTTGTAGCTTAGCAACTATTTTATCGTC 1837
```

RESULT 8
SPU40453
LOCUS SPU40453 4621 bp DNA linear PHG 11-MAR-1997
DEFINITION Streptococcus pyogenes phage T12 repressor, excisionase (xis), integrase (int) and erythrogenic toxin A precursor (speA) genes, complete cds.
ACCESSION U40453 M19350
VERSION U40453.1 GI:1877426
KEYWORDS erythrogenic toxin; type A streptococcal exotoxin.
SOURCE Streptococcus pyogenes phage T12
ORGANISM Streptococcus pyogenes phage T12
Viruses.
REFERENCE 1 (bases 2782 to 4621)
AUTHORS Weeks,C.R. and Ferretti,J.J.
TITLE Nucleotide sequence of the type A streptococcal exotoxin (erythrogenic toxin) gene from Streptococcus pyogenes bacteriophage T12
JOURNAL Infect. Immun. 52 (1), 144-150 (1986)
MEDLINE 86166804
PUBMED 3514452
REFERENCE 2 (bases 1 to 4621)
AUTHORS McShan,W.M., Tang,Y.F. and Ferretti,J.J.
TITLE Bacteriophage T12 of Streptococcus pyogenes integrates into the gene encoding a serine tRNA
JOURNAL Mol. Microbiol. 23 (4), 719-728 (1997)
MEDLINE 97206150
PUBMED 9157243
REFERENCE 3 (bases 1 to 4621)
AUTHORS McShan,W.M., Tang,Y.-F. and Ferretti,J.J.
TITLE Direct Submission
JOURNAL Submitted (09-NOV-1995) William M. McShan, Microbiology and Immunology, University of Oklahoma Health Sciences Center, 940 S. L. Young Blvd., Oklahoma City, OK 73104, USA
COMMENT On Mar 12, 1997 this sequence version replaced gi:216177.
FEATURES Location/Qualifiers
source 1. .4621
/organism="Streptococcus pyogenes phage T12"
/mol_type="genomic DNA"
/db_xref="taxon:35344"
/clone="p1179"
CDS 266. .1036
/note="lambda cI homolog"
/codon_start=1
/transl_table=11
/product="repressor"
/protein_id="AAC48865.1"
/db_xref="GI:1877427"
/translation="MDRFAEKSGLTKGYISMLEKNEHPKSKKPIIPTEETLLKVAKGM
GVDIDFVLSKLDSDQEIQINISPKNMLNMDNPSTPTTPKVELIPSTLQKINSTSSQLE
HKRQLNVLDYAETQLEQQNTGRKNHKPPTTPSNYYDHAASAGTGQYLNDVQVEKIELP
VDYDADFVIPVYGDSMEPKYHSGDYVFVKLSVELTDGDIGVFEYYGDAYIKQLLINDE
GAFLHSLNSKYEDILIDRDSDFRIIGEVVGSYSGNHSS"
gene 1532. .1750
/gene="xis"
CDS 1532. .1750
/gene="xis"
/codon_start=1
/transl_table=11
/product="excisionase"
/protein_id="AAC48866.1"
/db_xref="GI:1877428"
/translation="MFHGWNYNSYKANNNLLEIRHLEDIVKYNTIMRDQYKKGSNEYN
AFQQDIEQDKNQIALLFNEHQLLENIDN"
gene 1883. .2971
/gene="int"
CDS 1883. .2971
/gene="int"
/codon_start=1
/transl_table=11
/product="integrase"
/protein_id="AAC48867.1"
/db_xref="GI:1877429"
/translation="MKIKSYKKENGETAYKFVLYAGYVNGKRKYIRREGFKTKQAARE
TLISLQAELDKPKSSMTFGALTDQWLKEYEKTVQGSTYLKTERNINKHILPKLDKVKI
GDINPLLIQRLTEEWCNDLKYGGKILGLVRNILNLAVRYGYINNNPALPITPPKIKRK
RKMNNNFYTLDQLKQFLELVEKTDNIEKIALFRLLAFTGIRKGELLALTWDDLNGNTL
SINKAVTRTQVGLEIDVTKTKSSDRLISLDDETLEILQELHETFPTSTLMFQSKSGGI
MTPSLPRKWLLQIIKGIDLPQITIHGFRHTHASLLFESGLSLKQVQHRLGHGDLQTTM
NVYTHITQSAIDDIGTKFNQFVTNKQLD"
misc_feature 3014. .3109
/note="phage attachment site; 96 bp duplication with Streptococcus pyogenes chromosome at bacterial attachment site, 3' end of serine tRNA gene"
misc_feature 3014. .3015
/note="phage attachment site crossover point"
gene 3597. .4352
/gene="speA"
CDS 3597. .4352
/gene="speA"
/note="pyrogenic exotoxin A"
/codon_start=1
/transl_table=11
/product="erythrogenic toxin A precursor"
/protein_id="AAC48868.1"
/db_xref="GI:1877430"
/translation="MENNKKVLKKMVFFVLVTFLGLTISQEVFAQQDPDPSQLHRSSL
VKNLQNIYFLYEGDPVTHENVKSVDQLLSHDLIYNVSGPNYDKLKTELKNQEMATLFK
DKNVDIYGVEYYHLCYLCENAERSACIYGGVTNHEGNHLEIPKKIVVKVSIDGIQSLS
FDIETNKKMVTAQELDYKVRKYLTDNKQLYTNGPSKYETGYIKFIPKNKESFWFDFFP
EPEFTQSKYLMIYKDNETLDSNTSQIEVYLTTK"
sig_peptide 3597. .3686

```
                     /gene="speA"
                     /note="putative"
     mat_peptide     3687. .4349
                     /gene="speA"
                     /product="erythrogenic toxin A"
     terminator      4426. .4456
                     /note="putative transcription termination signal"
ORIGIN

  Query Match             90.1%;  Score 1668.4;  DB 7;  Length 4621;
  Best Local Similarity   97.4%;  Pred. No. 4e-230;
  Matches 1772;  Conservative    0;  Mismatches   36;  Indels   12;  Gaps    7;

Qy         40 AAGCTTACTTTTCGAATCAGGTCTATCCTTGAAACAGGTGCAACATAGATTAGGGCATGG 99
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2806 AAGCTTACTTTTCGAATCAGGTCTATCCTTGAAACAGGTGCAACATAGATTAGGGCATGG 2865

Qy        100 AGATTTACCAGACAACTATGAACGTATATACTCACATCACGCAATCGGCAATTGATGACA 159
              ||||||| ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2866 AGATTTA-CAGACAACTATGAACGTATATACTCACATCACGCAATCGGCAATTGATGACA 2924

Qy        160 TTGGAACTAAATTCAATCAATTTGTTACTAACAAGCAACTAGATTGACAACTAATTCTCA 219
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2925 TTGGAACTAAATTCAATCAATTTGTTACTAACAAGCAACTAGATTGACAACTAATTCTCA 2984

Qy        220 ACAAACGTTAATTTAACAACATTCAAGTAACTCCCACCAGCTCCATCAATGCTTACCGTA 279
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2985 ACAAACGTTAATTTAACAACATTCAAGTAACTCCCACCAGCTCCATCAATGCTTACCGTA 3044

Qy        280 AGTAATCATAACTTACTAAAACCTTGTTACATCAAGGTTTTTTCTTTTTGTCTTGTTCAT 339
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3045 AGTAATCATAACTTACTAAAACCTTGTTACATCAAGGTTTTTTCTTTTTGTCTTGTTCAT 3104

Qy        340 GAGTTACCATAACTTTCTATATTATTGACAACTAAATTGACAACTCTTCAATTATTTTTC 399
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3105 GAGTTACCATAACTTTCTATATTATTGACAACTAAATTGACAACTCTTCAATTATTTTTC 3164

Qy        400 TGTCTACTCAAAGTTTTCTTCATTTGATATAGTCTAATTCCACCATCACTTCTTCCACTC 459
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3165 TGTCTACTCAAAGTTTTCTTCATTTGATATAGTCTAATTCCACCATCACTTCTTCAGCCT 3224

Qy        460 TCTCTACCGTCACAACT--TCATCATCTCTCACTTTTTCGTGTGGTAACACATAATCAAA 517
              ||||  | |||||||||  ||||  ||    |  |  || || ||||||||||||||||
Db       3225 TCTCCTCTGTCACAACTTCTCATTTCCTTCGCTTCCTCATTGCGGCAACACATAATCAAA 3284

Qy        518 TATCTT-TCCGTTTTTACGCACTATCGCTACTGTGTCACCTAAAATATACCCCTT-ATCA 575
              |||||  |||||||||||||||  |||   |||||||  |||||||||||||||  |||
Db       3285 AATCTTGTCCGTTTTTACGCACGATCATAACTGTGTCTCCTAAAATATACCCCTTGCTCA 3344

Qy        576 ATCGCTTCTTTAAACTCATCTATATATAACATATTTCATCCTCCTACCTATCTATTCGTA 635
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3345 ATCGCTTCTTTAAACTCATCTATATATAACATATTTCATCCTCCTACCTATCTATTCGTA 3404

Qy        636 AAAAGATAAAAATAACTATTGTTTTTTTTGTTATTTTATAATAAAATTATTAATATAAGT 695
              |||||||||||||||||||||||| ||||||||||||||||||||||||||||||||||
Db       3405 AAAAGATAAAAATAACTATTGTTTTGTTTGTTATTTTATAATAAAATTATTAATATAAGT 3464

Qy        696 TAATGTTTTTTAAAAATATACAATTTTATTCTATTTATAGTTAGCTATTTTTTCATTGTT 755
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3465 TAATGTTTTTTAAAAATATACAATTTTATTCTATTTATAGTTAGCTATTTTTTCATTGTT 3524

Qy        756 AGTAATATTGGTGAATTGTAATAACCTTTTTAAATCTAGAGGAGAACCCAGATATAAAAT 815
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3525 AGTAATATTGGTGAATTGTAATAACCTTTTTAAATCTAGAGGAGAACCCAGATATAAAAT 3584

Qy        816 GGAGGAATATTAATGGAAAACAATAAAAAAGTATTGAAGAAAATGGTATTTTTTGTTTTA 875
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3585 GGAGGAATATTAATGGAAAACAATAAAAAAGTATTGAAGAAAATGGTATTTTTTGTTTTA 3644

Qy        876 GTGACATTTCTTGGACTAACAATCTCGCAAGAGGTATTTGCTCAACAAGACCCCGATCCA 935
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3645 GTGACATTTCTTGGACTAACAATCTCGCAAGAGGTATTTGCTCAACAAGACCCCGATCCA 3704

Qy        936 AGCCAACTTCACAGATCTAGTTTAGTTAAAAACCTTCAAAATATATATTTTCTTTATGAG 995
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3705 AGCCAACTTCACAGATCTAGTTTAGTTAAAAACCTTCAAAATATATATTTTCTTTATGAG 3764

Qy        996 GGTGACCCTGTTACTCACGAGAATGTGAAATCTGTTGATCAACTTTTATCTCACCATTTA 1055
              |||||||||||||||||||||||||||||||||||||||||||||||||||||| |||||
Db       3765 GGTGACCCTGTTACTCACGAGAATGTGAAATCTGTTGATCAACTTTTATCTCACGATTTA 3824

Qy       1056 ATATATAATGTTTCAGGGCCAAATTATGATAAATTAAAAACTGAACTTAAGAACCAAGAG 1115
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3825 ATATATAATGTTTCAGGGCCAAATTATGATAAATTAAAAACTGAACTTAAGAACCAAGAG 3884

Qy       1116 ATGGCAACTTTATTTAAGGATAAAAACGTTGATATTTATGGTGTAGAATATTACCATCTC 1175
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3885 ATGGCAACTTTATTTAAGGATAAAAACGTTGATATTTATGGTGTAGAATATTACCATCTC 3944

Qy       1176 TGTTATTTATGTGAAAATGCAGAAAGGAGTGCATGTATCTACGGAGGGGTAACAAATCAT 1235
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3945 TGTTATTTATGTGAAAATGCAGAAAGGAGTGCATGTATCTACGGAGGGGTAACAAATCAT 4004

Qy       1236 GAAGGGAATCATTTAGAAATTCCTAAAAAGATAGTCGTTAAAGTATCAATCGATGGTATC 1295
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       4005 GAAGGGAATCATTTAGAAATTCCTAAAAAGATAGTCGTTAAAGTATCAATCGATGGTATC 4064

Qy       1296 CAAAGCCTATCATTTGATATTGAAACAAATAAAAAAATGGTAACTGCTCAAGAATTAGAC 1355
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       4065 CAAAGCCTATCATTTGATATTGAAACAAATAAAAAAATGGTAACTGCTCAAGAATTAGAC 4124

Qy       1356 TATAAAGTTAGAAAATATCTTACAGATAATAAGCAACTATATACTAATGGACCTTCTAAA 1415
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       4125 TATAAAGTTAGAAAATATCTTACAGATAATAAGCAACTATATACTAATGGACCTTCTAAA 4184

Qy       1416 TATGAAACTGGATATATAAAGTTCATACCTAAGAATAAAGAAAGTTTTTGGTTTGATTTT 1475
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       4185 TATGAAACTGGATATATAAAGTTCATACCTAAGAATAAAGAAAGTTTTTGGTTTGATTTT 4244

Qy       1476 TTCCCTGAACCAGAATTTACTCAATCTAAATATCTTATGATATATAAAGATAATGAAACG 1535
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       4245 TTCCCTGAACCAGAATTTACTCAATCTAAATATCTTATGATATATAAAGATAATGAAACG 4304

Qy       1536 CTTGACTCAAACACAAGCCAAATTGAAGTCTACCTAACAACCAAGTAACTTTTTGCT--- 1592
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       4305 CTTGACTCAAACACAAGCCAAATTGAAGTCTACCTAACAACCAAGTAACTTTTTGCTTCG 4364

Qy       1593 -TTTGGCAACCTTACCTACTGCTGGATTTAGAAATTTTATTGCAATTCTTTTATTAATGT 1651
               |  ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       4365 GTACGGCAACCTTACCTACTGCTGGATTTAGAAATTTTATTGCAATTCTTTTATTAATGT 4424

Qy       1652 AAAAACCGCTCATTTGATGAGCGGTTTTGTCTTATCTAAAGGAGCTTTACCTCCTAATGC 1711
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       4425 AAAAACCGCTCATTTGATGAGCGGTTTTGTCTTATCTAAAGGAGCTTTACCTCCTAATGC 4484

Qy       1712 TGCAAAATTTTAAATGTTGGATTTTTGTATTTGTCTATTGTATTTGATGGGTAATCCCAT 1771
              |||||||||||||||||||||||||||||||||||||||||||||||||||| | ||||
Db       4485 TGCAAAATTTTAAATGTTGGATTTTTGTATTTGTCTATTGTATTTGATGG--TAATCCAT 4542

Qy       1772 TTTTCGACAGACATCGTCGTGCCACCTCTAACACCAAAATCATAGACAGGAGCTTGTAGC 1831
              |||||||||||||||||||| ||||||||||||||||||||||||||||||||||||||
Db       4543 TTTTCGACAGACATCGTCGTG-CACCTCTAACACCAAAATCATAGACAGGAGCTTGTAGC 4601

Qy       1832 TTAGCAACTATTTTATCGTC 1851
              ||||||||||||||||||||
Db       4602 TTAGCAACTATTTTATCGTC 4621


RESULT 9
AE014161
LOCUS       AE014161               57506 bp    DNA     linear   BCT 18-JUL-2002
DEFINITION  Streptococcus pyogenes MGAS315, section 26 of 37 of the complete
            genome.
ACCESSION   AE014161 AE014074
VERSION     AE014161.1  GI:21905010
KEYWORDS    .
```

```
SOURCE      Streptococcus pyogenes MGAS315
  ORGANISM  Streptococcus pyogenes MGAS315
            Bacteria; Firmicutes; Lactobacillales; Streptococcaceae;
            Streptococcus.
REFERENCE   1  (bases 1 to 57506)
  AUTHORS   Beres,S.B., Sylva,G.L., Barbian,K.D., Lei,B., Hoff,J.S.,
            Mammarella,N.D., Liu,M.-Y., Smoot,J.C., Porcella,S.F.,
            Parkins,L.D., McCormick,J.K., Leung,D.Y.M., Schlievert,P.M. and
            Musser,J.M.
  TITLE     Genome sequence of a serotype M3 strain of group A Streptococcus:
            Phage-encoded toxins, the high-virulence phenotype, and clone
            emergence
  JOURNAL   Proc. Natl. Acad. Sci. U.S.A. 99 (15), 10078-10083 (2002)
REFERENCE   2  (bases 1 to 57506)
  AUTHORS   Beres,S.B., Sylva,G.L., Barbian,K.D., Lei,B., Hoff,J.S.,
            Mammarella,N.D., Liu,M.-Y., Smoot,J.C., Porcella,S.F.,
            Parkins,L.D., McCormick,J.K., Leung,D.Y.M., Schlievert,P.M. and
            Musser,J.M.
  TITLE     Direct Submission
  JOURNAL   Submitted (14-JUN-2002) Laboratory of Human-Bacterial Pathogenesis,
            Rocky Mountain Laboratories, NIAID, NIH, 903 South Fourth St.,
            Hamilton, MT 59840, USA
FEATURES             Location/Qualifiers
     source          1. .38692
                     /organism="Streptococcus pyogenes MGAS315"
                     /mol_type="genomic DNA"
                     /strain="MGAS315"
                     /serotype="M3"
                     /db_xref="taxon:198466"
                     /focus
                     /note="group: A"
     source          complement(38693. .>57506)
                     /organism="Streptococcus pyogenes phage 315.5"
                     /mol_type="genomic DNA"
                     /db_xref="taxon:198542"
     gene            complement(141. .1073)
                     /gene="SpyM3_1270"
     CDS             complement(141. .1073)
                     /gene="SpyM3_1270"
                     /note="best non-GAS blastp hit: gb|AAK75827.1| (AE007467)
                     transporter, CorA family, putative [Streptococcus
                     pneumoniae TIGR4]"
                     /codon_start=1
                     /transl_table=11
                     /product="putative divalent cation transport protein"
                     /protein_id="AAM79877.1"
                     /db_xref="GI:21905011"
                     /translation="MPTLPKASSAITLINLDQLTEHDQEALVSEGLIDAEVFDYAKDK
                     NETSFMEENKEKLTIVFQILDRQEESYHPNHIPRVIPITLFLNDQSLYILGHDHSLTL
                     IEEVFPDLDESRSPRHLVFQLLTAFTKQYVPLMDEIAQQRDKLIEALRQKANKTNLES
                     LANLQSGTVYILMGSKQNEQMLAELKDMPGQQDLEEDEAEQLRDAIIEARQLSNMCDL
                     NTRVLKEISSSYNNVLSNNLNNNVTSLTIFSIGISIIAMVTSFYGMNVKLPFAKVDSV
                     WFWIVLTTSLVALLIMVVMYWYVHKRNRNASKRI"
     gene            complement(1165. .2124)
                     /gene="SpyM3_1271"
     CDS             complement(1165. .2124)
                     /gene="SpyM3_1271"
                     /note="best blastp hit: gb|AAL98140.1| (AE010071)
                     oxidoreductase [Streptococcus pyogenes MGAS8232]"
                     /codon_start=1
                     /transl_table=11
                     /product="putative oxidoreductase"
                     /protein_id="AAM79878.1"
                     /db_xref="GI:21905012"
                     /translation="MAKVRYGIVSTAQVAPRFIEGVRLAGNGEVVAVSSRSLDKAKAF
                     AAAHQLPKAYGSLDDMLLDASIDAIYVASINQDHFPAAKKALLAGKHVLVEKPFTLTS
                     AQAEELFALAQERGLFLMEAQKAVFIPMTQRLKTLITAGELGDIVSVSSTTAYPNIDH
                     VTWFRDLEKGGGTVHFMAPYALSYLPYLFDADIVSASGVAHFPKGQSDSQSKLVLALS
                     NGVLVDIFLTTHLSLPHELVIYGTKGRLEIPHFWKTTHATLVRPDGSRSVIEATMSSD
                     FEAEAHHVSQMILEGQLTSPVMTKSLTLSGVGLIEKLYRSWGK"
     gene            3134. .3913
                     /gene="SpyM3_1272"
     CDS             3134. .3913
                     /gene="SpyM3_1272"
                     /note="best blastp hit: gb|AAL98141.1| (AE010071)
                     hypothetical protein [Streptococcus pyogenes MGAS8232]"
                     /codon_start=1
                     /transl_table=11
                     /product="hypothetical protein"
                     /protein_id="AAM79879.1"
                     /db_xref="GI:21905013"
                     /translation="MLNSDFIISKSLANYIHHRRLEVGVSSTDLAEISNMSKSDWESF
                     EKNGGAIPLNSKDIILDLLFLERFPKEKECDFIDKLFEEAKENKLWPEKIYQTMGLTP
                     ALSFIAGCEILSDDINNDLEELSKLPKESHLGQLDTSLLLSLLPQQFITKYDYEFVYK
                     LSKVLAQYTSRNKVGSPYTAHSVIEEICLYLIAKESILYFESLDENSHLQLKELLDYN
                     DEWPFDIFDDMDSYTFLYTDIYIEEDSLYHFKNWFVPQFYL"
     gene            complement(4129. .6777)
                     /gene="valS"
                     /note="synonym: SpyM3_1273"
     CDS             complement(4129. .6777)
                     /gene="valS"
                     /note="best blastp hit: gb|AAL98142.1| (AE010071) putative
                     valine-tRNA ligase [Streptococcus pyogenes MGAS8232]"
                     /codon_start=1
                     /transl_table=11
                     /product="putative valine-tRNA ligase"
                     /protein_id="AAM79880.1"
                     /db_xref="GI:21905014"
                     /translation="MTELSPKYNPAEVEAGRYQKWLDADVFKPSGDQKAKPYSIVIPP
                     PNVTGKLHLGHAWDTTLQDIIIRQKRMQGFDTLWLPGMDHAGIATQAKVEERLRGQGI
                     TRYDLGREKFLDKVWEWKDEYATTIKEQWGKMGLSVDYSRERFTLDEGLSKAVRKVFV
                     DLYKKGWIYRGEFIINWDPAARTALSDIEVIHKDVEGAFYHMNYMLEDGSRALQVATT
                     RPETMFGDVAVAVNPEDPRYKDLIGKHVILPIVNKLIPIVGDEHADPEFGTGVVKITP
                     AHDPNDFEVGQRHNLPQVNVMNDDGTMNELAGDFAGMDRFEARQATVAKLEELGALVN
                     IEKRVHSVGHSERSGAVVEPRLSTQWFVKMDELAKQAMDNQETDDRVDFYPPRFNDTF
                     LQWMENVHDWVISRQLWWGHQIPAWYNAEGEIYVGEEAPEGDGWTQDEDVLDTWFSSA
                     LWPFSTMGWPDTDVEDFKRYFPTSTLVTGYDIIFFWVSRMIFQTLEFTGRQPFQNVLI
                     HGLIRDEEGRKMSKSLGNGIDPMDVIEKYGADSLRWFLSNGSAPGQDVRFSYEKMDAS
                     WNFINKIWNISRYILMNNEGLTLEEAESNVAKVAASEAGNVTDQWILHNLNETIAKVT
                     ENFDKFEFGVAGHILYNFIWEEFANWYVELTKEVLYSDNEAEKVITRSVLLYTLDKIL
                     RLLHPIMPFVTEEIYAQYAQGSIVTAAYPTVTPAFENEAAHKGVESLKDLIRAVRNAR
                     AEVNVAPSKPITILVKTADSELEDFFTSNVNYIKRFTNPEKLEISSAIAAPELAMTSI
                     ITGAEIYLPLADLLNVEEELARLDKELAKWQKELDMVGKKLGNERFVANAKPEVVQKE
                     KDKQADYQAKYDATQERIAEMQKLVK"
     gene            complement(6779. .7087)
                     /gene="SpyM3_1274"
     CDS             complement(6779. .7087)
                     /gene="SpyM3_1274"
                     /note="best non-GAS blastp hit: gb|AAK74723.1| (AE007366)
                     conserved domain protein [Streptococcus pneumoniae TIGR4]"
                     /codon_start=1
                     /transl_table=11
                     /product="conserved hypothetical protein"
                     /protein_id="AAM79881.1"
                     /db_xref="GI:21905015"
                     /translation="MLEKIQVVFQSYDQEVLFVELKTDIEERLKRNRTENRLKHKPLK
                     RNIEWSEQDIQSTMAYAVFNPEEPPKTLTHYQKINNTHLTAAETAQLIIQKMTHIKEN
                     "
     gene            complement(7112. .7294)
                     /gene="SpyM3_1275"
     CDS             complement(7112. .7294)
                     /gene="SpyM3_1275"
                     /note="best non-GAS blastp hit: gb|AAK74723.1| (AE007366)
                     conserved domain protein [Streptococcus pneumoniae TIGR4]"
                     /codon_start=1
                     /transl_table=11
                     /product="conserved hypothetical protein"
                     /protein_id="AAM79882.1"
                     /db_xref="GI:21905016"
                     /translation="MNLIIIGAQASGKMTIGQEVARQTGMTLFHNHDSIDFVLRFMPW
                     SQESTALIERIRFAFF"
     gene            complement(7291. .7419)
                     /gene="SpyM3_1276"
     CDS             complement(7291. .7419)
                     /gene="SpyM3_1276"
                     /note="best non-GAS blastp hit: gb|AAK99294.1| (AE008428)
```

```
                Conserved hypothetical protein [Streptococcus pneumoniae"
                /codon_start=1
                /transl_table=11
                /product="conserved hypothetical protein"
                /protein_id="AAM79883.1"
                /db_xref="GI:21905017"
                /translation="MACKLGFTLEANARDRKDVQGRRCGDMRFGLLRSEWEEQKQK"
     gene       complement(7895. .8290)
                /gene="SpyM3_1277"
     CDS        complement(7895. .8290)
                /gene="SpyM3_1277"
                /note="best non-GAS blastp hit: gb|AAK74720.1| (AE007366)
                hypothetical protein [Streptococcus pneumoniae TIGR4]"
                /codon_start=1
                /transl_table=11
                /product="conserved hypothetical protein"
                /protein_id="AAM79884.1"
                /db_xref="GI:21905018"
                /translation="MAKQKNRKKALKLEMKRNGLLKKAGQVFDKAVESVETAVDKTIS
                AGKNLVEKGSQTVENLTASKERLTLDEFAGQPVVAGIRSDLVETLYAEGIHSAQAFKE
                WTEKDLLALKGIGPATVKKLVENGASFKK"
     gene       complement(8311. .8565)
                /gene="SpyM3_1278"
     CDS        complement(8311. .8565)
                /gene="SpyM3_1278"
                /note="best non-GAS blastp hit: gb|AAK99291.1| (AE008428)
                hypothetical protein [Streptococcus pneumoniae R6]"
                /codon_start=1
                /transl_table=11
                /product="conserved hypothetical protein"
                /protein_id="AAM79885.1"
                /db_xref="GI:21905019"
                /translation="MSYEKEFLKDFEDWVKTQIQVNQLAMATSQEVAQEDGDERAKDA
                FIRYESKLDAYEFLLGKFDNYKNGKAFHDVPDELFGARHY"
     tRNA       8732. .8802
                /product="tRNA-Thr"
     gene       9044. .9796
                /gene="aroA.2"
                /note="synonym: SpyM3_1279"
     CDS        9044. .9796
                /gene="aroA.2"
                /note="best blastp hit: gb|AAK34362.1| (AE006589) putative
                chorismate mutase [Streptococcus pyogenes M1 GAS]"
                /codon_start=1
                /transl_table=11
                /product="putative chorismate mutase"
                /protein_id="AAM79886.1"
                /db_xref="GI:21905020"

 Query Match             76.0%;  Score 1406.6;  DB 1;  Length 57506;
 Best Local Similarity   99.3%;  Pred. No. 7.5e-193;
 Matches 1423;  Conservative    0;  Mismatches    9;  Indels    1;  Gaps    1;

Qy      419 TCATTTGATATAGTCTAATTCCACCATCACTTCTTCCACTCTCTCTACCGTCACAACTTC 478
            || |||    | ||||||||||||||||||||||||||||||||||||||||||||||
Db    38775 TCTTTTTTATTTGTCTAATTCCACCATCACTTCTTCCACTCTCTCTACCGTCACAACTTC 38834

Qy      479 ATCATCTCTCACTTTTTCGTGTGGTAACACATAATCAAATATCTTTCCGTTTTTACGCAC 538
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    38835 ATCATCTCTCACTTTTTCGTGTGGTAACACATAATCAAATATCTTTCCGTTTTTACGCAC 38894

Qy      539 TATCGCTACTGTGTCACCTAAAATATACCCCTTATCAATCGCTTCTTTAAACTCATCTAT 598
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    38895 TATCGCTACTGTGTCACCTAAAATATACCCCTTATCAATCGCTTCTTTAAACTCATCTAT 38954

Qy      599 ATATAACATATTTCATCCTCCTACCTATCTATTCGTAAAAAGATAAAAATAACTATTGTT 658
            |||||||||||||||||||||||||||||||||||||||||||||||||||||||||| |
Db    38955 ATATAACATATTTCATCCTCCTACCTATCTATTCGTAAAAAGATAAAAATAACTATTG-T 39013

Qy      659 TTTTTTGTTATTTTATAATAAAATTATTAATATAAGTTAATGTTTTTTAAAAATATACAA 718
            |||||||||||||||||||||||||||||||||||||||||||| ||||||||||||
Db    39014 TTTTTTGTTATTTTATAATAAAATTATTAATATAAGTTAATGTTTTTTTAAAATATACAA 39073

Qy      719 TTTTATTCTATTTATAGTTAGCTATTTTTTCATTGTTAGTAATATTGGTGAATTGTAATA 778
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    39074 TTTTATTCTATTTATAGTTAGCTATTTTTTCATTGTTAGTAATATTGGTGAATTGTAATA 39133

Qy      779 ACCTTTTTAAATCTAGAGGAGAACCCAGATATAAAATGGAGGAATATTAATGGAAAACAA 838
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    39134 ACCTTTTTAAATCTAGAGGAGAACCCAGATATAAAATGGAGGAATATTAATGGAAAACAA 39193

Qy      839 TAAAAAAGTATTGAAGAAAATGGTATTTTTTGTTTTAGTGACATTTCTTGGACTAACAAT 898
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    39194 TAAAAAAGTATTGAAGAAAATGGTATTTTTTGTTTTAGTGACATTTCTTGGACTAACAAT 39253

Qy      899 CTCGCAAGAGGTATTTGCTCAACAAGACCCCGATCCAAGCCAACTTCACAGATCTAGTTT 958
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    39254 CTCGCAAGAGGTATTTGCTCAACAAGACCCCGATCCAAGCCAACTTCACAGATCTAGTTT 39313

Qy      959 AGTTAAAAACCTTCAAAATATATATTTTCTTTATGAGGGTGACCCTGTTACTCACGAGAA 1018
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    39314 AGTTAAAAACCTTCAAAATATATATTTTCTTTATGAGGGTGACCCTGTTACTCACGAGAA 39373

Qy     1019 TGTGAAATCTGTTGATCAACTTTTATCTCACCATTTAATATATAATGTTTCAGGGCCAAA 1078
            |||||||||||||||||||||||||||||| ||||||||||||||||||||||||||||
Db    39374 TGTGAAATCTGTTGATCAACTTTTATCTCACGATTTAATATATAATGTTTCAGGGCCAAA 39433

Qy     1079 TTATGATAAATTAAAAACTGAACTTAAGAACCAAGAGATGGCAACTTTATTTAAGGATAA 1138
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    39434 TTATGATAAATTAAAAACTGAACTTAAGAACCAAGAGATGGCAACTTTATTTAAGGATAA 39493

Qy     1139 AAACGTTGATATTTATGGTGTAGAATATTACCATCTCTGTTATTTATGTGAAAATGCAGA 1198
            |||| ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    39494 AAACATTGATATTTATGGTGTAGAATATTACCATCTCTGTTATTTATGTGAAAATGCAGA 39553

Qy     1199 AAGGAGTGCATGTATCTACGGAGGGGTAACAAATCATGAAGGGAATCATTTAGAAATTCC 1258
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    39554 AAGGAGTGCATGTATCTACGGAGGGGTAACAAATCATGAAGGGAATCATTTAGAAATTCC 39613

Qy     1259 TAAAAAGATAGTCGTTAAAGTATCAATCGATGGTATCCAAAGCCTATCATTTGATATTGA 1318
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    39614 TAAAAAGATAGTCGTTAAAGTATCAATCGATGGTATCCAAAGCCTATCATTTGATATTGA 39673

Qy     1319 AACAAATAAAAAAAATGGTAACTGCTCAAGAATTAGACTATAAAGTTAGAAAATATCTTAC 1378
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    39674 AACAAATAAAAAAATGGTAACTGCTCAAGAATTAGACTATAAAGTTAGAAAATATCTTAC 39733

Qy     1379 AGATAATAAGCAACTATATACTAATGGACCTTCTAAATATGAAACTGGATATATAAAGTT 1438
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    39734 AGATAATAAGCAACTATATACTAATGGACCTTCTAAATATGAAACTGGATATATAAAGTT 39793

Qy     1439 CATACCTAAGAATAAAGAAAGTTTTTGGTTTGATTTTTTCCCTGAACCAGAATTTACTCA 1498
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    39794 CATACCTAAGAATAAAGAAAGTTTTTGGTTTGATTTTTTCCCTGAACCAGAATTTACTCA 39853

Qy     1499 ATCTAAATATCTTATGATATATAAAGATAATGAAACGCTTGACTCAAACACAAGCCAAAT 1558
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    39854 ATCTAAATATCTTATGATATATAAAGATAATGAAACGCTTGACTCAAACACAAGCCAAAT 39913

Qy     1559 TGAAGTCTACCTAACAACCAAGTAACTTTTTGCTTTTGGCAACCTTACCTACTGCTGGAT 1618
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    39914 TGAAGTCTACCTAACAACCAAGTAACTTTTTGCTTTTGGCAACCTTACCTACTGCTGGAT 39973

Qy     1619 TTAGAAATTTTATTGCAATTCTTTTATTAATGTAAAAACCGCTCATTTGATGAGCGGTTT 1678
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    39974 TTAGAAATTTTATTGCAATTCTTTTATTAATGTAAAAACCGCTCATTTGATGAGCGGTTT 40033

Qy     1679 TGTCTTATCTAAAGGAGCTTTACCTCCTAATGCTGCAAAATTTTAAATGTTGGATTTTTG 1738
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    40034 TGTCTTATCTAAAGGAGCTTTACCTCCTAATGCTGCAAAATTTTAAATGTTGGATTTTTG 40093

Qy     1739 TATTTGTCTATTGTATTTGATGGGTAATCCCATTTTTCGACAGACATCGTCGTGCCACCT 1798
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    40094 TATTTGTCTATTGTATTTGATGGGTAATCCCATTTTTCGACAGACATCGTCGTGCCACCT 40153

Qy     1799 CTAACACCAAAATCATAGACAGGAGCTTGTAGCTTAGCAACTATTTTATCGTC 1851
```

```
                     ||||||||||||||||||||||||||||||||||||||||||||||||||
Db      40154 CTAACACCAAAATCATAGACAGGAGCTTGTAGCTTAGCAACTATTTTATCGTC 40206


RESULT 10
AP005142/c
LOCUS       AP005142               310850 bp    DNA     linear   BCT 13-JUN-2003
DEFINITION  Streptococcus pyogenes SSI-1 DNA, complete genome, section 2/6.
ACCESSION   AP005142 BA000034
VERSION     AP005142.1  GI:28810440
KEYWORDS    .
SOURCE      Streptococcus pyogenes SSI-1
  ORGANISM  Streptococcus pyogenes SSI-1
            Bacteria; Firmicutes; Lactobacillales; Streptococcaceae;
            Streptococcus.
REFERENCE   1
  AUTHORS   Nakagawa,I., Kurokawa,K., Yamashita,A., Nakata,M., Tomiyasu,Y.,
            Okahashi,N., Kawabata,S., Yamazaki,K., Shiba,T., Yasunaga,T.,
            Hayashi,H., Hattori,M. and Hamada,S.
  TITLE     Genome Sequence of an M3 Strain of Streptococcus pyogenes Reveals a
            Large-Scale Genomic Rearrangement in Invasive Strains and New
            Insights into Phage Evolution
  JOURNAL   Genome Res. 13 (6), 1042-1055 (2003)
  MEDLINE   22683278
   PUBMED   12799345
REFERENCE   2  (bases 1 to 310850)
  AUTHORS   Yamashita,A., Nakagawa,I., Kurokawa,K., Nakata,M., Tomiyasu,Y.,
            Yamazaki,K., Okahashi,N., Kawabata,S., Yasunaga,T., Hattori,M.,
            Hayashi,H. and Hamada,S.
  TITLE     Direct Submission
  JOURNAL   Submitted (01-MAY-2002) Ken Kurokawa, Osaka University, Genome
            Information Research Center; 3-1, Yamadaoka, Suita, Osaka 565-0871,
            Japan (E-mail:ken@gen-info.osaka-u.ac.jp,
            URL:http://www.gen-info.osaka-u.ac.jp/, Tel:81-6-6879-8365,
            Fax:81-6-6879-2047)
COMMENT     genome project
            This clone was isolated from a patient presenting with toxic shock
            like syndrome.
FEATURES             Location/Qualifiers
     source          1. .310850
                     /organism="Streptococcus pyogenes SSI-1"
                     /mol_type="genomic DNA"
                     /strain="SSI-1"
                     /db_xref="taxon:193567"
     gene            complement(89. .1243)
                     /gene="SPs0283"
     CDS             complement(89. .1243)
                     /gene="SPs0283"
                     /note="similar to GB:AAL98401.1 (AE010096) percent
                     identity 100 in 374 aa"
                     /codon_start=1
                     /evidence=not_experimental
                     /transl_table=11
                     /product="putative A/G-specific adenine glycosylase"
                     /protein_id="BAC63378.1"
                     /db_xref="GI:28810441"
                     /translation="MIELKDYGINMWDNETIASFRRTLLGWYDQEKRDLPWRRTTNPY
                     YIWVSEIMLQQTQVNTVIPYYKRFLEWFPQIKDLADAPEEQLLKAWEGLGYYSRVRNM
                     QKAAQQVMVDFGGIFPHTYDDIASLKGIGPYTAGAIASISFNLPEPAVDGNVMRVMAR
                     LFEVNYDIGDPKNRKIFQAIMEILIDPDRPGDFNQALMDLGTDIESAKTPRPDESPIR
                     FFNAAYLNGTYGKYPIKNPKKKPKPMRIQAFVIRNQNGQYLLEKNTKGRLLGGFWSFP
                     IIETSPLSQQLDLFDDNQSNPIIWQTQNETFEREYQLKPQWTDNHFPNIKHTFSHQKW
                     TIELIEGVVKATDLPNAPHLKWVAIEDFSLYPFATPQKKMLETYLKQKNA"
     gene            1411. .1704
                     /gene="SPs0284"
     CDS             1411. .1704
                     /gene="SPs0284"
                     /note="similar to GB:AAK34554.1 (AE006609) percent
                     identity 100 in 97 aa"
                     /codon_start=1
                     /evidence=not_experimental
                     /transl_table=11
                     /product="hypothetical protein"
                     /protein_id="BAC63379.1"
                     /db_xref="GI:28810442"
                     /translation="MFQKLKNLIDRYSWFVFIMMILVLFLNYIYPFDPSLSGLQLEWL
                     IMTQVFPGLGIVCAMVAIIFKKFNLFVLALILYFALWLKLFLAFSFFTYFLGV"
     gene            1877. .2167
                     /gene="SPs0285"
     CDS             1877. .2167
                     /gene="SPs0285"
                     /note="similar to GB:AAK34553.1 (AE006609) percent
                     identity 100 in 96 aa"
                     /codon_start=1
                     /evidence=not_experimental
                     /transl_table=11
                     /product="30S ribosomal protein S6"
                     /protein_id="BAC63380.1"
                     /db_xref="GI:28810443"
                     /translation="MAKYEILYIIRPNIEEEAKNALVARFDSILTDNGATVVESKDWE
                     KRRLAYEINDFREGLYHIVNLEATDAAALNEFDRLSKINGDILRHMIVKLDA"
     gene            2189. .2680
                     /gene="SPs0286"
     CDS             2189. .2680
                     /gene="SPs0286"
                     /note="similar to GB:AAK34552.1 (AE006609) percent
                     identity 100 in 163 aa"
                     /codon_start=1
                     /evidence=not_experimental
                     /transl_table=11
                     /product="putative single strand binding protein (phage
                     associated)"
                     /protein_id="BAC63381.1"
                     /db_xref="GI:28810444"
                     /translation="MINNVVLVGRMTKDAELRYTPSQVAVATFTLAVNRTFKSQNGER
                     EADFINCVIWRQPAENLANWAKKGALIGVTGRIQTRNYENQQGQRVYVTEVVADNFQM
                     LESRATREGGSTGSFNGGFNNNTSSSNSYSAPAQQTPNFGRDDSPFGNSNPMDISDDD
                     LPF"
     gene            2845. .3084
                     /gene="SPs0287"
     CDS             2845. .3084
                     /gene="SPs0287"
                     /note="similar to GB:AAK34551.1 (AE006609) percent
                     identity 100 in 79 aa"
                     /codon_start=1
                     /evidence=not_experimental
                     /transl_table=11
                     /product="30S ribosomal protein S18"
                     /protein_id="BAC63382.1"
                     /db_xref="GI:28810445"
                     /translation="MAQQRRGGFKRRKKVDFIAANKIEYVDYKDTELLSRFVSERGKI
                     LPRRVTGTSAKNQRKVTTAIKRARVMALMPYVNED"
     gene            complement(3217. .3873)
                     /gene="SPs0288"
     CDS             complement(3217. .3873)
                     /gene="SPs0288"
                     /note="similar to GB:AAL98396.1 (AE010095) percent
                     identity 100 in 218 aa"
                     /codon_start=1
                     /evidence=not_experimental
                     /transl_table=11
                     /product="hypothetical protein"
                     /protein_id="BAC63383.1"
                     /db_xref="GI:28810446"
                     /translation="MELQELTKKNQEFIHTATNKLIQDGKSDEDIKLILEEAIPAILE
                     NQKKGVTARNLLGTPTAWAASFSQDPSQRAAETDKNTNPWLMWLDTSLLFIGIVALLN
                     GIMTFFNTNATVTGLISLLALGFGGGASMYATYYFIYRHLGKDKSLRPSWFKIIAALS
                     LAMLIWIALYSATAFLPTSLNPQLPPLALLIIGGVSLALRYYLQRKYNIQNTMSPVNK
                     "
     gene            complement(4005. .4949)
                     /gene="SPs0289"
     CDS             complement(4005. .4949)
                     /gene="SPs0289"
                     /note="similar to GB:AAK34549.1 (AE006609) percent
                     identity 100 in 314 aa"
                     /codon_start=1
```

```
                     /evidence=not_experimental
                     /transl_table=11
                     /product="putative divalent cation transport protein"
                     /protein_id="BAC63384.1"
                     /db_xref="GI:28810447"
                     /translation="MKQMFLSSAIEFKEIETFEPGAWIKLVNPSQEESMKIADQFNID
                     ISDLRAPLDVEETSRIAVEDDYTLIIVDVPIYEERNNKSYYITMPLGIIVTENAVITT
                     CLHDMTLFDHFHNRRVKNFYTFMKTRFVFQILYRNAELFLTALRTIDRQSERLEAQLE
                     AATRNEELIDMMELEKSIVYLKASLKFNERIVKKLSSSTSSLKKYIEDEDLLEDTLIE
                     TQQAIEMAGIYENVLNAMTETTASIINNNQNTIMKTLALMTMALDIPTVIFSAYGMNF
                     QNNWLPLNGLEHAFWYITLIAMLLSSFVVIYFIRKKWF"
     gene            5122. .7980
                     /gene="SPs0290"
     CDS             5122. .7980
                     /gene="SPs0290"
                     /note="similar to GB:AAL98394.1 (AE010095) percent
                     identity 99 in 942 aa"
                     /codon_start=1
                     /evidence=not_experimental
                     /transl_table=11
                     /product="putative excinuclease ABC subunit A"
                     /protein_id="BAC63385.1"
                     /db_xref="GI:28810448"
                     /translation="MRRVPLGTFKMQNKIIIHGARAHNLKNIDVEIPRDKLVVVTGLS
                     GSGKSSLAFDTIYAEGQRRYVESLSAYARQFLGNMEKPDVDSIDGLSPAISIDQKTTS
                     KNPRSTVGTVTEINDYLRLLYARVGTPYCINGHGAITASSAEQIVEQVLALPERTRMQ
                     ILAPIVRRKKGQHKTIFEKIQKDGYVRVRVDGDIFDVTEVPELSKSKMHNIEVVIDRL
                     VNKDGIRSRLFDSVEAALRLGDGYLMIDTMDGNELLFSEHYSCPVCGFTVPELEPRLF
                     SFNAPFGSCPTCDGLGIKLEVDLDLVVPDPSKSLKEGALAPWNPISSNYYPTMLEQAM
                     ASFGVDMDTPFEALTEEERDLVLYGSGDREFHFHYVNDFGGERNIDIPFEGVVTNVNR
                     RYHETNSDYTRNVMRGYMNELTCATCHGYRLNDQALCVHVGGEEGPHIGQISELSIAD
                     HLQLLEELELTENESTIAKPIVKEIHDRLTFLNNVGLNYLTLSRAAGTLSGGESQRIR
                     LATQIGSNLSGVLYILDEPSIGLHQRDNDRLIESLKKMRDLGNTLIVVEHDEDTMMQA
                     DWLIDVGPGAGEFGGEIIASGTPKQVAKNKKSITGQYLSGKKFIPVPLERRSGNGRFI
                     EIKGAAQNNLQSLDVRFPLGKFIAVTGVSGSGKSTLVNSILKKAVAQKLNRNADKPGK
                     YHSISGIEHIERLIDIDQSPIGRTPRSNPATYTGVFDDIRDLFAQTNEAKIRGYKKGR
                     FSFNVKGGRCEACSGDGIIKIEMHFLPDVYVPCEVCHGRRYNSETLEVHYKEKNIAEV
                     LDMTVDDALVFFSAIPKIARKIQTIKDVGLGYVTLGQPATTLSGGEAQRMKLASELHK
                     RSTGKSLYILDEPTTGLHTDDIARLLKVLERFVDDGNTVLVIEHNLDVIKSADHIIDL
                     GPEGGVGGGQIVATGTPEEVAQVKESYTGHYLKVKLQQ"
     gene            8096. .9169
                     /gene="SPs0291"
     CDS             8096. .9169
                     /gene="SPs0291"
                     /note="similar to GB:AAL98393.1 (AE010095) percent
                     identity 98 in 357 aa"
                     /codon_start=1
                     /evidence=not_experimental
                     /transl_table=11
                     /product="putative aminopeptidase P, XAA-pro
                     aminopeptidase"
                     /protein_id="BAC63386.1"
                     /db_xref="GI:28810449"
                     /translation="MSGFLEQRLGHCLRQMVEKGLEALLVTHLTNSYYLTGFSGTAAT
                     VLITAKRRVLITDSRYTLLAKASVEGFDIIESRTPLKVVAELLEADQIDCLGFEDQVS
                     FSFYQAMQAELSGITLLAQSGFVEHLRLIKDASEIETIAKACSISDKAFEDALDFIKP
                     GTTTERDLANFLNFRMRQYGASGTSFDIIVASGYRSAMPHGRASDKVIQNKESLTMDF
                     GCYYNHYVSDMTRTIHIGQVTDEEREIYALVLAANKALIAKASAGMTYSDFDGIPRQL
                     ITEAGYGSHFTHGIGHGIGLDIHENPFFGKSEQLLQAGMVVTDEPGIYLDNKYGVRIE
                     DDLVITKTGCQVLTLAPKELIVL"

  Query Match          76.0%;  Score 1406.6;  DB 1;  Length 310850;
  Best Local Similarity  99.3%;  Pred. No. 5.1e-193;
  Matches 1423;  Conservative    0;  Mismatches    9;  Indels    1;  Gaps    1;

Qy        419 TCATTTGATATAGTCTAATTCCACCATCACTTCTTCCACTCTCTCTACCGTCACAACTTC 478
              || |||    | ||||||||||||||||||||||||||||||||||||||||||||||
Db     272059 TCTTTTTTATTTGTCTAATTCCACCATCACTTCTTCCACTCTCTCTACCGTCACAACTTC 272000

Qy        479 ATCATCTCTCACTTTTTCGTGTGGTAACACATAATCAAATATCTTTCCGTTTTTACGCAC 538
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     271999 ATCATCTCTCACTTTTTCGTGTGGTAACACATAATCAAATATCTTTCCGTTTTTACGCAC 271940

Qy        539 TATCGCTACTGTGTCACCTAAAATATACCCCTTATCAATCGCTTCTTTAAACTCATCTAT 598
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     271939 TATCGCTACTGTGTCACCTAAAATATACCCCTTATCAATCGCTTCTTTAAACTCATCTAT 271880

Qy        599 ATATAACATATTTCATCCTCCTACCTATCTATTCGTAAAAAGATAAAAATAACTATTGTT 658
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||| |
Db     271879 ATATAACATATTTCATCCTCCTACCTATCTATTCGTAAAAAGATAAAAATAACTATTG-T 271821

Qy        659 TTTTTTGTTATTTTATAATAAAATTATTAATATAAGTTAATGTTTTTTAAAAATATACAA 718
              |||||||||||||||||||||||||||||||||||||||||||||||| ||||||||||
Db     271820 TTTTTTGTTATTTTATAATAAAATTATTAATATAAGTTAATGTTTTTTTAAAATATACAA 271761

Qy        719 TTTTATTCTATTTATAGTTAGCTATTTTTTCATTGTTAGTAATATTGGTGAATTGTAATA 778
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     271760 TTTTATTCTATTTATAGTTAGCTATTTTTTCATTGTTAGTAATATTGGTGAATTGTAATA 271701

Qy        779 ACCTTTTTAAATCTAGAGGAGAACCCAGATATAAAATGGAGGAATATTAATGGAAAACAA 838
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     271700 ACCTTTTTAAATCTAGAGGAGAACCCAGATATAAAATGGAGGAATATTAATGGAAAACAA 271641

Qy        839 TAAAAAAGTATTGAAGAAAATGGTATTTTTTGTTTTAGTGACATTTCTTGGACTAACAAT 898
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     271640 TAAAAAAGTATTGAAGAAAATGGTATTTTTTGTTTTAGTGACATTTCTTGGACTAACAAT 271581

Qy        899 CTCGCAAGAGGTATTTGCTCAACAAGACCCCGATCCAAGCCAACTTCACAGATCTAGTTT 958
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     271580 CTCGCAAGAGGTATTTGCTCAACAAGACCCCGATCCAAGCCAACTTCACAGATCTAGTTT 271521

Qy        959 AGTTAAAAACCTTCAAAATATATATTTTCTTTATGAGGGTGACCCTGTTACTCACGAGAA 1018
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     271520 AGTTAAAAACCTTCAAAATATATATTTTCTTTATGAGGGTGACCCTGTTACTCACGAGAA 271461

Qy       1019 TGTGAAATCTGTTGATCAACTTTTATCTCACCATTTAATATATAATGTTTCAGGGCCAAA 1078
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     271460 TGTGAAATCTGTTGATCAACTTTTATCTCACGATTTAATATATAATGTTTCAGGGCCAAA 271401

Qy       1079 TTATGATAAATTAAAAACTGAACTTAAGAACCAAGAGATGGCAACTTTATTTAAGGATAA 1138
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     271400 TTATGATAAATTAAAAACTGAACTTAAGAACCAAGAGATGGCAACTTTATTTAAGGATAA 271341

Qy       1139 AAACGTTGATATTTATGGTGTAGAATATTACCATCTCTGTTATTTATGTGAAAATGCAGA 1198
              |||| ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     271340 AAACATTGATATTTATGGTGTAGAATATTACCATCTCTGTTATTTATGTGAAAATGCAGA 271281

Qy       1199 AAGGAGTGCATGTATCTACGGAGGGGTAACAAATCATGAAGGGAATCATTTAGAAATTCC 1258
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     271280 AAGGAGTGCATGTATCTACGGAGGGGTAACAAATCATGAAGGGAATCATTTAGAAATTCC 271221

Qy       1259 TAAAAAGATAGTCGTTAAAGTATCAATCGATGGTATCCAAAGCCTATCATTTGATATTGA 1318
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     271220 TAAAAAGATAGTCGTTAAAGTATCAATCGATGGTATCCAAAGCCTATCATTTGATATTGA 271161

Qy       1319 AACAAATAAAAAAAATGGTAACTGCTCAAGAATTAGACTATAAAGTTAGAAAATATCTTAC 1378
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     271160 AACAAATAAAAAAATGGTAACTGCTCAAGAATTAGACTATAAAGTTAGAAAATATCTTAC 271101

Qy       1379 AGATAATAAGCAACTATATACTAATGGACCTTCTAAATATGAAACTGGATATATAAAGTT 1438
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     271100 AGATAATAAGCAACTATATACTAATGGACCTTCTAAATATGAAACTGGATATATAAAGTT 271041

Qy       1439 CATACCTAAGAATAAAGAAAGTTTTTGGTTTGATTTTTTCCCTGAACCAGAATTTACTCA 1498
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     271040 CATACCTAAGAATAAAGAAAGTTTTTGGTTTGATTTTTTCCCTGAACCAGAATTTACTCA 270981

Qy       1499 ATCTAAATATCTTATGATATATAAAGATAATGAAACGCTTGACTCAAACACAAGCCAAAT 1558
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     270980 ATCTAAATATCTTATGATATATAAAGATAATGAAACGCTTGACTCAAACACAAGCCAAAT 270921

Qy       1559 TGAAGTCTACCTAACAACCAAGTAACTTTTTGCTTTTGGCAACCTTACCTACTGCTGGAT 1618
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     270920 TGAAGTCTACCTAACAACCAAGTAACTTTTTGCTTTTGGCAACCTTACCTACTGCTGGAT 270861

Qy       1619 TTAGAAATTTTATTGCAATTCTTTTATTAATGTAAAAACCGCTCATTTGATGAGCGGTTT 1678
```

```
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      270860 TTAGAAATTTTATTGCAATTCTTTTATTAATGTAAAAACCGCTCATTTGATGAGCGGTTT 270801

Qy        1679 TGTCTTATCTAAAGGAGCTTTACCTCCTAATGCTGCAAAATTTTAAATGTTGGATTTTTG 1738
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      270800 TGTCTTATCTAAAGGAGCTTTACCTCCTAATGCTGCAAAATTTTAAATGTTGGATTTTTG 270741

Qy        1739 TATTTGTCTATTGTATTTGATGGGTAATCCCATTTTTCGACAGACATCGTCGTGCCACCT 1798
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      270740 TATTTGTCTATTGTATTTGATGGGTAATCCCATTTTTCGACAGACATCGTCGTGCCACCT 270681

Qy        1799 CTAACACCAAAATCATAGACAGGAGCTTGTAGCTTAGCAACTATTTTATCGTC 1851
                |||||||||||||||||||||||||||||||||||||||||||||||||||
Db      270680 CTAACACCAAAATCATAGACAGGAGCTTGTAGCTTAGCAACTATTTTATCGTC 270628


RESULT 11
AE009982/c
LOCUS       AE009982               11900 bp    DNA     linear   BCT 03-APR-2002
DEFINITION  Streptococcus pyogenes strain MGAS8232, section 30 of 173 of the
            complete genome.
ACCESSION   AE009982 AE009949
VERSION     AE009982.1  GI:19747622
KEYWORDS    .
SOURCE      Streptococcus pyogenes MGAS8232
  ORGANISM  Streptococcus pyogenes MGAS8232
            Bacteria; Firmicutes; Lactobacillales; Streptococcaceae;
            Streptococcus.
REFERENCE   1  (bases 1 to 11900)
  AUTHORS   Smoot,J.C., Barbian,K.D., Van Gompel,J.J., Smoot,L.M.,
            Chaussee,M.S., Sylva,G.L., Sturdevant,D.E., Ricklefs,S.M.,
            Porcella,S.F., Parkins,L.D., Beres,S.B., Campbell,D.S., Smith,T.M.,
            Zhang,Q., Kapur,V., Daly,J.A., Veasy,L.G. and Musser,J.M.
  TITLE     Genome sequence and comparative microarray analysis of serotype M18
            group A Streptococcus strains associated with acute rheumatic fever
            outbreaks
  JOURNAL   Proc. Natl. Acad. Sci. U.S.A. 99 (7), 4668-4673 (2002)
  MEDLINE   21927593
   PUBMED   11917108
REFERENCE   2  (bases 1 to 11900)
  AUTHORS   Smoot,J.C., Barbian,K.D., Van Gompel,J.J., Smoot,L.M.,
            Chaussee,M.S., Sylva,G.L., Sturdevant,D.E., Ricklefs,S.M.,
            Porcella,S.F., Parkins,L.D., Beres,S.B., Campbell,D.S., Smith,T.M.,
            Zhang,Q., Kapur,V., Daly,J.A., Veasy,L.George. and Musser,J.M.
  TITLE     Direct Submission
  JOURNAL   Submitted (31-JAN-2002) Laboratory of Human Bacterial
            Pathogenesis/Rocky Mountain Laboratories/NIAID/NIH, 903 S. 4th St.,
            Hamilton, MT 59840, USA
FEATURES             Location/Qualifiers
     source          1. .11900
                     /organism="Streptococcus pyogenes MGAS8232"
                     /mol_type="genomic DNA"
                     /strain="MGAS8232"
                     /db_xref="taxon:186103"
     misc_feature    <1. .2669
                     /note="phi speA"
     gene            90. .1307
                     /gene="spyM18_0391"
     CDS             90. .1307
                     /gene="spyM18_0391"
                     /note="best blastp match emb|CAB07986.1| (Z93946)
                     N-acetylmuramoyl-L-alanine amidase [bacteriophage Dp-1]"
                     /codon_start=1
                     /transl_table=11
                     /product="putative N-acetylmuramoyl-L-alanine amidase"
                     /protein_id="AAL97140.1"
                     /db_xref="GI:19747623"
                     /translation="MVVDTEKAIAWMGLKKGRVSYSMDCRNGPDSYDCSSAICSALIY
                     AGASNPGWLLNTEYMHDWLVQNGFELIADNEDWDSQRADIAIWGFRGQSAGAGGHVVM
                     FIDADNIIHCNYANNNITIDNYNQTAAASGWMYSYVYRYTGAKAQPATNKSIDELVQE
                     VLSGKHGSGEQRKISLGANYDAVQAKVNELLKQPQVAEQSPAVKQDGDLLFNGAVLKK
                     AILDKILAKCKEHDILPSYAITVLHFEGLWGQSAVGRADNNWGGMTWTGQGNRPSGIT
                     VTQGTERPAVEGGHYMHYASVDDFLTDWFYLLRADGSYRVSGAKTFSEAVKGMFKVGG
                     ATYDYAASGFDSYIVGMAGRLKAIEQENGSLAKYDQQADIDVGQSDKIDVVIDSLEIN
                     INGVTYTATKKPI"
     gene            complement(1429. .2184)
                     /gene="speA"
                     /note="synonym: spyM18_0393"
     CDS             complement(1429. .2184)
                     /gene="speA"
                     /note="best blastp match sp|P08095|SPEA_STRPY EXOTOXIN
                     TYPE A PRECURSOR (SCARLET FEVER TOXIN) (ERYTHROGENIC
                     TOXIN) (SPE A)"
                     /codon_start=1
                     /transl_table=11
                     /product="exotoxin type A precursor"
                     /protein_id="AAL97141.1"
                     /db_xref="GI:19747624"
                     /translation="MENNKKVLKKMVFFVLVTFLGLTISQEVFAQQDPDPSQLHRSSL
                     VKNLQNIYFLYEGDPVTHENVKSVDQLLSHDLIYNVSGPNYDKLKTELKNQEMATLFK
                     DKNVDIYGVEYYHLCYLCENAERSACIYGGVTNHEGNHLEIPKKIVVKVSIDGIQSLS
                     FDIETNKKMVTAQELDYKVRKYLTDNKQLYTNGPSKYETGYIKFIPKNKESFWFDFFP
                     EPEFTQSKYLMIYKDNETLDSNTSQIEVYLTTK"
     gene            2404. .2586
                     /gene="spyM18_0394"
     CDS             2404. .2586
                     /gene="spyM18_0394"
                     /note="best blastp match hypothetical"
                     /codon_start=1
                     /transl_table=11
                     /product="hypothetical phage protein"
                     /protein_id="AAL97142.1"
                     /db_xref="GI:19747625"
                     /translation="MLYIDEFKEAIDKGYILGDTVAIVRKNGKIFDYVLPGEGVRPWE
                     VLTEEKVEEVMRELDK"
     misc_feature    2662. .2681
                     /note="phi SpeA attR"
     gene            2895. .3497
                     /gene="spyM18_0396"
     CDS             2895. .3497
                     /gene="spyM18_0396"
                     /note="best blastp match gb|AAK33395.1| (AE006499)
                     hypothetical protein [Streptococcus pyogenes M1 GAS]"
                     /codon_start=1
                     /transl_table=11
                     /product="hypothetical protein"
                     /protein_id="AAL97143.1"
                     /db_xref="GI:19747626"
                     /translation="MVKNQFPLVADGIAISDPAKQMALYENEDLITNIRGYYQDKEYD
                     DIARNEEFTAKATSRQAPSSKRFCSNDEKHHYVKEARQKAKQDLKEKRQAYLAKEMTY
                     VPKQVSKKQQPADSSPSQKQATTEMSRFTKKLHQDNYILAELPKEYKEPKNLPQQGTT
                     KKNNYDFLKSSQIYNNKEMRQQREKTIAQELNLSRFEDLP"
     gene            3537. .4865
                     /gene="spyM18_0398"
     CDS             3537. .4865
                     /gene="spyM18_0398"
                     /note="best blastp match gb|AAK33396.1| (AE006499)
                     putative UDP-N-acetyl muramate-alanine ligase
                     [Streptococcus pyogenes M1 GAS]"
                     /codon_start=1
                     /transl_table=11
                     /product="putative UDP-N-acetyl muramate-alanine ligase"
                     /protein_id="AAL97144.1"
                     /db_xref="GI:19747627"
                     /translation="MSKTYHFIGIKGSGMSALALMLHQMGHMVQGSDVEKYYFTQRGL
                     EQAGITILPFSEDNITPDMELIVGNAFRENNKEVAYALRHQIPFKRYHDFLGDFMKSF
                     ISFAVAGAHGKTSTTGLLSHVLKNITDTSYLIGDGTGRGSANAQYFVFESDEYERHFM
                     PYHPEYSIITNIDFDHPDYFTGIADVRNAFNDYAKQVKKALFVYGEDDELKKIEAPAP
                     IYYYGFEEGNDFIAYDITRTTNGSDFKVKHQGEVIGQFHVPAYGKHNILNATAVIANL
                     FVAGIDMALVADHLKTFSGVKRRFTEKIVNDTIIIDDFAHHPTEIVATIDAARQKYPS
                     KEIVAIFQPHTFTRTIALLEDFACALNEADSVYLAQIYGSAREVDKGEVKVEDLAAKI
                     IKPSQVVTVENVSPLLDHDNAVYVFMGAGDIQLYEHSFEELLANLTKNNQ"
     gene            4911. .5393
                     /gene="spyM18_0399"
     CDS             4911. .5393
                     /gene="spyM18_0399"
```

```
                /note="best blastp match gb|AAK33397.1| (AE006499)
                putative arylalkylamine n-acetyltransferase [Streptococcus
                pyogenes M1 GAS]"
                /codon_start=1
                /transl_table=11
                /product="putative arylalkylamine n-acetyltransferase"
                /protein_id="AAL97145.1"
                /db_xref="GI:19747628"
                /translation="MLIRQVQGSDLEVIATIESDNFSPQEATTRAVLEERIRLIPDTF
                LVALIDQEIVGYIEGPVVTTPILEDSLFHGVTKNPKTGGYIAITSLSIAKHFQQQGVG
                TALLAALKDLVVAQQRTGLILTCHDYLISYYEMNGFINQGISESQHGGTLWYQMIWKL
                "
gene            5511. .7079
                /gene="spyM18_0401"
CDS             5511. .7079
                /gene="spyM18_0401"
                /note="best blastp match gb|AAK33398.1| (AE006499)
                putative aminodeoxychorismate lyase [Streptococcus
                pyogenes M1 GAS]"
                /codon_start=1
                /transl_table=11
                /product="putative aminodeoxychorismate lyase"
                /protein_id="AAL97146.1"
                /db_xref="GI:19747629"
                /translation="MTDFKDKDQQDQQRSFKEQILAELEKANQIRKEKEEELFQKELE
                AKEAARRTAQLYAEYKRQDAFQKESIAHNNKTAKHFQAIKDAVMTSEAIKPTLLSEKE
                NSSLKTTSKRVVQANELQETASKESQVPLTIEKGHSVRRKLSKRQQTERAAKKISTVL
                ISSIIITLLAVTLAGAGYVYSALNPVDKNSDAFVQVEIPSGSGNKLIGQILQKKGLIK
                NSTVPSFYTKFKNFTNFQSGYYNLQKSMSLEEIASALQEGGTAEPTKPSLGKILIPEG
                YTIKQIAKAVEHNSKGKTKKAKTPFNEKDFLDLVTDEAFIQDMVKRYPKLLATIPTKE
                KAIYRLEGYLFPATYNYYKETTMRELVEDMLAAMDATLVPYYDKIAASGKTVNEVLTL
                ASLVEKEGSTDDDRRQIASVFYNRLNSGMALQSNIAILYAMGKLGEKTTLAEDATIDT
                TINSPYNIYTNTGLMPGPVASSGVSAIEATLNPASTDYLYFVANVHTGEVYYAKTFEE
                HSANVEKYVNSQIQ"
gene            7104. .7634
                /gene="spyM18_0402"
CDS             7104. .7634
                /gene="spyM18_0402"
                /note="best blastp match gb|AAK33399.1| (AE006499)
                putative transcription elongation factor [Streptococcus
                pyogenes M1 GAS]"
                /codon_start=1
                /transl_table=11
                /product="putative transcription elongation factor"
                /protein_id="AAL97147.1"
                /db_xref="GI:19747630"
                /translation="MFVFLIRSKKKREKTIMAEKTYPMTLTEKEQLEKELEELKLVRR
                PEIVERIKIARSYGDLSENSEYDAAKDEQAFVEGQISTLETKIRYAEIIDSDAVAKDE
                VAIGKTVIVQEVGTTDKDTYHIVGAAGADIFSGKISNESPIAQALIGKKTGDKVRIES
                PAATYDVEIISVEKTN"
gene            complement(8258. .9181)
                /gene="spyM18_0404"
CDS             complement(8258. .9181)
                /gene="spyM18_0404"
                /note="best blastp match gb|AAK33400.1| (AE006499)
                hypothetical protein [Streptococcus pyogenes M1 GAS]"
                /codon_start=1
                /transl_table=11
                /product="hypothetical protein"
                /protein_id="AAL97148.1"
                /db_xref="GI:19747631"
                /translation="MKLTLNRILFSGLALSILFTLTGCVGRDAHGNPKGMIWEFLGKP
                MSYFIDYFANNAGLGYGLAIIIVTIIVRTLILPLGLYQSWKASYQSEKMTFLKPVFEP
                INKRIKQASSQEEKMAAQTELMAAQRAHGINPLGGIGCLPLLIQMPFFSAMYFAAQYT
                KGVSTSTFMGIDLGSRSLVLTAIIAALYFFQSWLSMMAVSEEQREQMKTMMYTMPIMM
                IFMSFSLPAGVGLYWLVGGFFSIIQQLITTYLLKPRLHKQIKEEYAKTPPKAYQSTSS
                RKDVTPSQNMEQANLPKKIKSNRNAGKQRKR"
gene            complement(9263. .9541)
                /gene="spyM18_0405"
CDS             complement(9263. .9541)
                /gene="spyM18_0405"
                /note="best blastp match gb|AAK33401.1| (AE006499)
```

```
Query Match           66.2%;  Score 1225.4;  DB 1;  Length 11900;
Best Local Similarity 97.5%;  Pred. No. 1e-166;
Matches 1255;  Conservative    0;  Mismatches   31;  Indels    1;  Gaps    1;

Qy       422 TTTGATATAGTCTAATTCCACCATCACTTCTTCCACTCTCTCTACCGTCACAACTTCATC 481
Db      2589 TTTTTATTTGTCTAATTCCCTCATCACTTCTTCCACTTTTTCCTCTGTCAACACCTCCCA 2530

Qy       482 ATCTCTCACTTTTTCGTGTGGTAACACATAATCAAATATCTTTCCGTTTTTACGCACTAT 541
Db      2529 TGGCCTGACTCCCTCTCCTGGTAACACATAATCAAATATCTTTCCGTTTTTACGCACTAT 2470

Qy       542 CGCTACTGTGTCACCTAAAATATACCCCTTATCAATCGCTTCTTTAAACTCATCTATATA 601
Db      2469 CGCTACTGTGTCACCTAAAATATACCCCTTATCAATCGCTTCTTTAAACTCATCTATATA 2410

Qy       602 TAACATATTTCATCCTCCTACCTATCTATTCGTAAAAAGATAAAAATAACTATTGTTTTT 661
Db      2409 TAACATATTTCATCCTCCTACCTATCTATTCGTAAAAAGATAAAAATAACTATTG-TTTT 2351

Qy       662 TTTGTTATTTTATAATAAAATTATTAATATAAGTTAATGTTTTTTAAAAATATACAATTT 721
Db      2350 TTTGTTATTTTATAATAAAATTATTAATATAAGTTAATGTTTTTCAAAAATATACAATTT 2291

Qy       722 TATTCTATTTATAGTTAGCTATTTTTTCATTGTTAGTAATATTGGTGAATTGTAATAACC 781
Db      2290 TATTCTATTTATAGTTAGCTATTTTTTCATTGTTAGTAATATTGGTGAATTGTAATAACC 2231

Qy       782 TTTTTAAATCTAGAGGAGAACCCAGATATAAAATGGAGGAATATTAATGGAAAACAATAA 841
Db      2230 TTTTTAAATCTAGAGGAGAACCCAGATATAAAATGGAGGAATATTAATGGAAAACAATAA 2171

Qy       842 AAAAGTATTGAAGAAAATGGTATTTTTTGTTTTAGTGACATTTCTTGGACTAACAATCTC 901
Db      2170 AAAAGTATTGAAGAAAATGGTATTTTTTGTTTTAGTGACATTTCTTGGACTAACAATCTC 2111

Qy       902 GCAAGAGGTATTTGCTCAACAAGACCCCGATCCAAGCCAACTTCACAGATCTAGTTTAGT 961
Db      2110 GCAAGAGGTATTTGCTCAACAAGACCCCGATCCAAGCCAACTTCACAGATCTAGTTTAGT 2051

Qy       962 TAAAAACCTTCAAAATATATATTTTCTTTATGAGGGTGACCCTGTTACTCACGAGAATGT 1021
Db      2050 TAAAAACCTTCAAAATATATATTTTCTTTATGAGGGTGACCCTGTTACTCACGAGAATGT 1991

Qy      1022 GAAATCTGTTGATCAACTTTTATCTCACCATTTAATATATAATGTTTCAGGGCCAAATTA 1081
Db      1990 GAAATCTGTTGATCAACTTTTATCTCACGATTTAATATATAATGTTTCAGGGCCAAATTA 1931

Qy      1082 TGATAAATTAAAAACTGAACTTAAGAACCAAGAGATGGCAACTTTATTTAAGGATAAAAA 1141
Db      1930 TGATAAATTAAAAACTGAACTTAAGAACCAAGAGATGGCAACTTTATTTAAGGATAAAAA 1871

Qy      1142 CGTTGATATTTATGGTGTAGAATATTACCATCTCTGTTATTTATGTGAAAATGCAGAAAG 1201
Db      1870 CGTTGATATTTATGGTGTAGAATATTACCATCTCTGTTATTTATGTGAAAATGCAGAAAG 1811

Qy      1202 GAGTGCATGTATCTACGGAGGGGTAACAAATCATGAAGGGAATCATTTAGAAATTCCTAA 1261
Db      1810 GAGTGCATGTATCTACGGAGGGGTAACAAATCATGAAGGGAATCATTTAGAAATTCCTAA 1751

Qy      1262 AAAGATAGTCGTTAAAGTATCAATCGATGGTATCCAAAGCCTATCATTTGATATTGAAAC 1321
Db      1750 AAAGATAGTCGTTAAAGTATCAATCGATGGTATCCAAAGCCTATCATTTGATATTGAAAC 1691

Qy      1322 AAATAAAAAAAATGGTAACTGCTCAAGAATTAGACTATAAAGTTAGAAAATATCTTACAGA 1381
Db      1690 AAATAAAAAAAATGGTAACTGCTCAAGAATTAGACTATAAAGTTAGAAAATATCTTACAGA 1631

Qy      1382 TAATAAGCAACTATATACTAATGGACCTTCTAAATATGAAACTGGATATATAAAGTTCAT 1441
Db      1630 TAATAAGCAACTATATACTAATGGACCTTCTAAATATGAAACTGGATATATAAAGTTCAT 1571

Qy      1442 ACCTAAGAATAAAGAAAGTTTTTGGTTTGATTTTTTCCCTGAACCAGAATTTACTCAATC 1501
```

```
Db      1570 ACCTAAGAATAAAGAAAGTTTTTGGTTTGATTTTTTCCCTGAACCAGAATTTACTCAATC 1511

Qy      1502 TAAATATCTTATGATATATAAAGATAATGAAACGCTTGACTCAAACACAAGCCAAATTGA 1561
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      1510 TAAATATCTTATGATATATAAAGATAATGAAACGCTTGACTCAAACACAAGCCAAATTGA 1451

Qy      1562 AGTCTACCTAACAACCAAGTAACTTTTTGCTTTTGGCAACCTTACCTACTGCTGGATTTA 1621
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      1450 AGTCTACCTAACAACCAAGTAACTTTTTGCTTTTGGCAACCTTACCTACTGCTGGATTTA 1391

Qy      1622 GAAATTTTATTGCAATTCTTTTATTAATGTAAAAACCGCTCATTTGATGAGCGGTTTTGT 1681
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      1390 GAAATTTTATTGCAATTCTTTTATTAATGTAAAAACCGCTCATTTGATGAGCGGTTTTGT 1331

Qy      1682 CTTATCTAAAGGAGCTTTACCTCCTAA 1708
             ||||||||||||||||||||||||||
Db      1330 CTTATCTAAAGGAGCTTTACCTCCTAA 1304
```

```
RESULT 12
SPSPEA
LOCUS       SPSPEA                  1031 bp    DNA     linear   BCT 21-MAR-1995
DEFINITION  Streptococcus pyrogenes speA gene for exotoxin type A (scarlet
            fever toxin).
ACCESSION   X03929
VERSION     X03929.1  GI:47441
KEYWORDS    exotoxin; exotoxin type A; speA gene.
SOURCE      Streptococcus pyogenes
  ORGANISM  Streptococcus pyogenes
            Bacteria; Firmicutes; Lactobacillales; Streptococcaceae;
            Streptococcus.
REFERENCE   1  (bases 1 to 1031)
  AUTHORS   Johnson,L.P., L'Italien,J.J. and Schlievert,P.M.
  TITLE     Streptococcal pyrogenic exotoxin type A (scarlet fever toxin) is
            related to Staphylococcus aureus enterotoxin B
  JOURNAL   Mol. Gen. Genet. 203 (2), 354-356 (1986)
  MEDLINE   86284313
   PUBMED   3526093
FEATURES             Location/Qualifiers
     source          1. .1031
                     /organism="Streptococcus pyogenes"
                     /mol_type="genomic DNA"
                     /db_xref="taxon:1314"
     RBS             67. .71
                     /note="pot. rRNA binding site"
     CDS             79. .831
                     /note="unnamed protein product; precursor (aa -30 to 220)"
                     /codon_start=1
                     /transl_table=11
                     /protein_id="CAA27568.1"
                     /db_xref="GI:47442"
                     /db_xref="GOA:P08095"
                     /db_xref="SWISS-PROT:P08095"
                     /translation="MENNKEVLKKMVFFVLMKFLGLTILPKGICSTRPKPSQLQRSNL
                     VKTFKIYIFFMRVTLVTHENVKSVDQLLSHDLIYNVSGPNYDKLKTELKNQEMATLFK
                     DKNVDIYGVEYYHLCYLCENAERSACLYGGVTNHEGNHLEIPKKIVVKVSIDGIQSLS
                     FDIEQIKNGNCSRISYTVRKYLTDNKQLYTNGPSKYETGYIKFIPKNKESFWFDFFPE
                     PEFTQSKYLMIYKDNETLDSNTSQIEVYLTTK"
     sig_peptide     79. .168
                     /note="signal peptide (aa -30 to -1)"
     mat_peptide     169. .828
                     /product="mature exotoxin type A (aa 1-220)"
     misc_signal     904. .920
                     /note="palindrome pot. transcription terminator"
     misc_signal     1003. .1016
                     /note="palindrome pot. transcription terminator"
ORIGIN

  Query Match             48.2%;  Score 892.8;  DB 1;  Length 1031;
  Best Local Similarity   95.7%;  Pred. No. 8.7e-119;
  Matches  993;  Conservative    0;  Mismatches   37;  Indels    8;  Gaps    7;
```

```
Qy       750 ATTGTTAGTAATATTGGTGAATTGTAATAACCTTTTTAAATCTAGAGGAGAACCCAGATA 809
             ||||||| ||||| |||| |||||||||||||||||||||||| ||||||||||||||||
Db         1 ATTGTTAATAATAATGGTTAATTGTAATAACCTTTTTAAATCTAAAGGAGAACCCAGATA 60

Qy       810 TAAAATGGAGGAATATTAATGGAAAACAATAAAAAAGTATTGAAGAAAATGGTATTTTTT 869
             ||||||||||||||||||||||||||||||||||| ||||||||||||||||||||||||
Db        61 TAAAATGGAGGAATATTAATGGAAAACAATAAAGAAGTATTGAAGAAAATGGTATTTTTT 120

Qy       870 GTTTTAGTGACATTTCTTGGACTAACAATCTCGCAAGAGGTATTTGCTCAACAAGACCCC 929
             |||||| ||| |||||||||||||||||| || | | |  || | | | |||
Db       121 GTTTTAATGAAATTTCTTGGACTAACAATCTTGCCAAAAGGTATTTGTTCAACAAGACCC 180

Qy       930 GATCCAAGCCAACTTCACAGATCTAGTTTAGTTAAAAACCTTCAAAATATATATTTTCTT 989
             | || |||||| | || ||||||| ||||||| |||||||||||||||||||||||||||
Db       181 AAGCCCAGCCAATTACAAAGATCTAATTTAGTT-AAAACCTTCAAAATATATATTTTCTT 239

Qy       990 TATGAGGGTGACCCT-GTTACTCACGAGAATGTGAAATCTGTTGATCAACTTTTATCTCA 1048
             |||||||||||||| |||||||||||||||||||||||||||||||||||||| ||
Db       240 TATGAGGGTGACCCTGGTTACTCACGAGAATGTGAAATCTGTTGATCAACTTTTATCACA 299

Qy      1049 CCATTTAATATATAATGTTTCAGGGCCAAATTATGATAAATTAAAAACTGAACTTAAGAA 1108
             | ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       300 CGATTTAATATATAATGTTTCAGGGCCAAATTATGATAAATTAAAAACTGAACTTAAGAA 359

Qy      1109 CCAAGAGATGGCAACTTTATTTAAGGATAAAAACGTTGATATTTATGGTGTAGAATATTA 1168
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       360 CCAAGAGATGGCAACTTTATTTAAGGATAAAAACGTTGATATTTATGGTGTAGAATATTA 419

Qy      1169 CCATCTCTGTTATTTATGTGAAAATGCAGAAAGGAGTGCATGTATCTACGGAGGGGTAAC 1228
             ||||||||||||||||||||||||||||||||||||||||||||| ||||||||||||||
Db       420 CCATCTCTGTTATTTATGTGAAAATGCAGAAAGGAGTGCATGTCTCTACGGAGGGGTAAC 479

Qy      1229 AAATCATGAAGGGAATCATTTAGAAATTCCTAAAAAGATAGTCGTTAAAGTATCAATCGA 1288
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       480 AAATCATGAAGGGAATCATTTAGAAATTCCTAAAAAGATAGTCGTTAAAGTATCAATCGA 539

Qy      1289 TGGTATCCAAAGCCTATCATTTGATATTGAAACAAATAAAAAAATGGTAACTGCTCAAGA 1348
             |||||||||||||||||||||||||||||||  ||  ||||||||||||||||||||||
Db       540 TGGTATCCAAAGCCTATCATTTGATATTGAA--CAAATAAAAAATGGTAACTGCTCAAGA 597

Qy      1349 ATTAGACTATAAAGTTAGAAAATATCTTACAGATAATAAGCAACTATATACTAATGGACC 1408
             ||||| ||||| |||||||||||||||||||||||||||||||||||||||||||||||
Db       598 ATTAG-CTATACAGTTAGAAAATATCTTACAGATAATAAGCAACTATATACTAATGGACC 656

Qy      1409 TTCTAAATATGAAACTGGATATATAAAGTTCATACCTAAGAATAAAGAAAGTTTTTGGTT 1468
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       657 TTCTAAATATGAAACTGGATATATAAAGTTCATACCTAAGAATAAAGAAAGTTTTTGGTT 716

Qy      1469 TGATTTTTTCCCTGAACCAGAATTTACTCAATCTAAATATCTTATGATATATAAAGATAA 1528
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       717 TGATTTTTTCCCTGAACCAGAATTTACTCAATCTAAATATCTTATGATATATAAAGATAA 776

Qy      1529 TGAAACGCTTGACTCAAACACAAGCCAAATTGAAGTCTACCTAACAACCAAGTAACTTTT 1588
             |||||||||||||||||||||||||||||||||||||||||||||||||||||| |||
Db       777 TGAAACGCTTGACTCAAACACAAGCCAAATTGAAGTCTACCTAACAACCAAGTAAC-TTT 835

Qy      1589 TGCTTTTGGCAACCTTACCTACTGCTGGATTTAGAAATTTTATTGCAATTCTTTTATTAA 1648
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       836 TGCTTTTGGCAACCTTACCTACTGCTGGATTTAGAAATTTTATTGCAATTCTTTTATTAA 895

Qy      1649 TGTAAAAACCGCTCATTTGATGAGCGGTTTTGTCTTATCTAAAGGAGCTTTACCTCCTAA 1708
             |||||||| ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       896 TGTAAAAA-CGCTCATTTGATGAGCGGTTTTGTCTTATCTAAAGGAGCTTTACCTCCTAA 954

Qy      1709 TGCTGCAAAATTTTAAATGTTGGATTTTTGTATTTGTCTATTGTATTTGATGGGTAATCC 1768
             |||||||||||||||||||||||||||||||||||| ||||||||||||||||||||||
Db       955 TGCTGCAAAATTTTAAATGTTGGATTTTTGTATTTGTC-ATTGTATTTGATGGGTAATCC 1013

Qy      1769 CATTTTTCGACAGACATC 1786
             ||||||||||||||||||
Db      1014 CATTTTTCGACAGACATC 1031
```

```
RESULT 13
AX110260
LOCUS       AX110260                 756 bp    DNA     linear   PAT 29-MAY-2002
DEFINITION  Sequence 993 from Patent WO0123604.
ACCESSION   AX110260
VERSION     AX110260.1  GI:13926552
KEYWORDS    .
SOURCE      Streptococcus pyogenes
  ORGANISM  Streptococcus pyogenes
            Bacteria; Firmicutes; Lactobacillales; Streptococcaceae;
            Streptococcus.
REFERENCE   1
  AUTHORS   Bergeron,M.G., Boissinot,M., Huletsky,A., m Nard,C., Ouellette,M.,
            Picard,F.J. and Roy,P.H.
  TITLE     Highly conserved genes and their use to generate probes and primers
            for detection of microorganisms
  JOURNAL   Patent: WO 0123604-A 993 05-APR-2001;
            Infectio Diagnostic (I.D.I.) INC. (CA)
FEATURES             Location/Qualifiers
     source          1. .756
                     /organism="Streptococcus pyogenes"
                     /mol_type="unassigned DNA"
                     /db_xref="taxon:1314"
ORIGIN

  Query Match             40.8%;  Score 754.4;  DB 6;  Length 756;
  Best Local Similarity   99.9%;  Pred. No. 6.5e-99;
  Matches  755;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy        828 ATGGAAAACAATAAAAAAGTATTGAAGAAAATGGTATTTTTTGTTTTAGTGACATTTCTT 887
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 ATGGAAAACAATAAAAAAGTATTGAAGAAAATGGTATTTTTTGTTTTAGTGACATTTCTT 60

Qy        888 GGACTAACAATCTCGCAAGAGGTATTTGCTCAACAAGACCCCGATCCAAGCCAACTTCAC 947
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 GGACTAACAATCTCGCAAGAGGTATTTGCTCAACAAGACCCCGATCCAAGCCAACTTCAC 120

Qy        948 AGATCTAGTTTAGTTAAAAACCTTCAAAATATATATTTTCTTTATGAGGGTGACCCTGTT 1007
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 AGATCTAGTTTAGTTAAAAACCTTCAAAATATATATTTTCTTTATGAGGGTGACCCTGTT 180

Qy       1008 ACTCACGAGAATGTGAAATCTGTTGATCAACTTTTATCTCACCATTTAATATATAATGTT 1067
              |||||||||||||||||||||||||||||||||||| ||||||||||||||||||||||
Db        181 ACTCACGAGAATGTGAAATCTGTTGATCAACTTTTATCTCACGATTTAATATATAATGTT 240

Qy       1068 TCAGGGCCAAATTATGATAAATTAAAAACTGAACTTAAGAACCAAGAGATGGCAACTTTA 1127
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 TCAGGGCCAAATTATGATAAATTAAAAACTGAACTTAAGAACCAAGAGATGGCAACTTTA 300

Qy       1128 TTTAAGGATAAAAACGTTGATATTTATGGTGTAGAATATTACCATCTCTGTTATTTATGT 1187
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        301 TTTAAGGATAAAAACGTTGATATTTATGGTGTAGAATATTACCATCTCTGTTATTTATGT 360

Qy       1188 GAAAATGCAGAAAGGAGTGCATGTATCTACGGAGGGGTAACAAATCATGAAGGGAATCAT 1247
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        361 GAAAATGCAGAAAGGAGTGCATGTATCTACGGAGGGGTAACAAATCATGAAGGGAATCAT 420

Qy       1248 TTAGAAATTCCTAAAAAGATAGTCGTTAAAGTATCAATCGATGGTATCCAAAGCCTATCA 1307
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        421 TTAGAAATTCCTAAAAAGATAGTCGTTAAAGTATCAATCGATGGTATCCAAAGCCTATCA 480

Qy       1308 TTTGATATTGAAACAAATAAAAAAATGGTAACTGCTCAAGAATTAGACTATAAAGTTAGA 1367
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        481 TTTGATATTGAAACAAATAAAAAAATGGTAACTGCTCAAGAATTAGACTATAAAGTTAGA 540

Qy       1368 AAATATCTTACAGATAATAAGCAACTATATACTAATGGACCTTCTAAATATGAAACTGGA 1427
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        541 AAATATCTTACAGATAATAAGCAACTATATACTAATGGACCTTCTAAATATGAAACTGGA 600

Qy       1428 TATATAAAGTTCATACCTAAGAATAAAGAAAGTTTTTGGTTTGATTTTTTCCCTGAACCA 1487
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        601 TATATAAAGTTCATACCTAAGAATAAAGAAAGTTTTTGGTTTGATTTTTTCCCTGAACCA 660

Qy       1488 GAATTTACTCAATCTAAATATCTTATGATATATAAAGATAATGAAACGCTTGACTCAAAC 1547
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        661 GAATTTACTCAATCTAAATATCTTATGATATATAAAGATAATGAAACGCTTGACTCAAAC 720

Qy       1548 ACAAGCCAAATTGAAGTCTACCTAACAACCAAGTAA 1583
              ||||||||||||||||||||||||||||||||||||
Db        721 ACAAGCCAAATTGAAGTCTACCTAACAACCAAGTAA 756


RESULT 14
SP156SPEA
LOCUS       SP156SPEA                708 bp    DNA     linear   BCT 23-OCT-1991
DEFINITION  S.pyogenes strain MGAS156 speA gene (allele 1) for type A exotoxin.
ACCESSION   X61560
VERSION     X61560.1  GI:47287
KEYWORDS    exotoxin; exotoxin type A.
SOURCE      Streptococcus pyogenes
  ORGANISM  Streptococcus pyogenes
            Bacteria; Firmicutes; Lactobacillales; Streptococcaceae;
            Streptococcus.
REFERENCE   1
  AUTHORS   Nelson,K., Schlievert,P.M., Selander,R.K. and Musser,J.M.
  TITLE     Characterization and clonal distribution of four alleles of the
            speA gene encoding pyrogenic exotoxin A (scarlet fever toxin) in
            Streptococcus pyogenes
  JOURNAL   J. Exp. Med. 174 (5), 1271-1274 (1991)
  MEDLINE   92044323
   PUBMED   1940804
REFERENCE   2  (bases 1 to 708)
  AUTHORS   Nelson,K.
  TITLE     Direct Submission
  JOURNAL   Submitted (01-SEP-1991) K. Nelson, Pennsylvania State University,
            Inst of Mol Evolutionary Genetics, University Park, PA 16802, USA
FEATURES             Location/Qualifiers
     source          1. .708
                     /organism="Streptococcus pyogenes"
                     /mol_type="genomic DNA"
                     /strain="MGAS156"
                     /isolate="Nebraska"
                     /db_xref="taxon:1314"
                     /chromosome="bacteriophage encoded"
                     /clone="direct PCR fragment"
     gene            1. .708
                     /gene="speA"
     CDS             <1. .>708
                     /gene="speA"
                     /codon_start=1
                     /transl_table=11
                     /product="type A exotoxin"
                     /protein_id="CAA43758.1"
                     /db_xref="GI:47288"
                     /db_xref="GOA:P97163"
                     /db_xref="SPTREMBL:P97163"
                     /translation="KKMVFFVLVTFLGLTISQEVFAQQDPDPSQLHRSSLVKNLQNIY
                     FLYEGDPVTHENVKSVDQLLSHDLIYNVSGPNYDKLKTELKNQEMATLFKDKNVDIYG
                     VEYYHLCYLCENAERSACIYGGVTNHEGNHLEIPKKIVVKVSIDGIQSLSFDIETNKK
                     MVTAQELDYKVRKYLTDNKQLYTNGPSKYETGYIKFIPKNKESFWFDFFPEPEFTQSK
                     YLMIYKDNETLDSNTSQI"
     sig_peptide     <1. .66
                     /gene="speA"
     mat_peptide     67. .>708
                     /gene="speA"
                     /product="type A exotoxin"
ORIGIN

  Query Match             38.2%;  Score 706.4;  DB 1;  Length 708;
  Best Local Similarity   99.9%;  Pred. No. 5e-92;
  Matches  707;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy        852 AAGAAAATGGTATTTTTTGTTTTAGTGACATTTCTTGGACTAACAATCTCGCAAGAGGTA 911
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 AAGAAAATGGTATTTTTTGTTTTAGTGACATTTCTTGGACTAACAATCTCGCAAGAGGTA 60
```

```
Qy        912 TTTGCTCAACAAGACCCCGATCCAAGCCAACTTCACAGATCTAGTTTAGTTAAAAACCTT 971
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 TTTGCTCAACAAGACCCCGATCCAAGCCAACTTCACAGATCTAGTTTAGTTAAAAACCTT 120

Qy        972 CAAAATATATATTTTCTTTATGAGGGTGACCCTGTTACTCACGAGAATGTGAAATCTGTT 1031
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 CAAAATATATATTTTCTTTATGAGGGTGACCCTGTTACTCACGAGAATGTGAAATCTGTT 180

Qy       1032 GATCAACTTTTATCTCACCATTTAATATATAATGTTTCAGGGCCAAATTATGATAAATTA 1091
              |||||||||||||||| ||||||||||||||||||||||||||||||||||||||||||
Db        181 GATCAACTTTTATCTCACGATTTAATATATAATGTTTCAGGGCCAAATTATGATAAATTA 240

Qy       1092 AAAACTGAACTTAAGAACCAAGAGATGGCAACTTTATTTAAGGATAAAAACGTTGATATT 1151
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 AAAACTGAACTTAAGAACCAAGAGATGGCAACTTTATTTAAGGATAAAAACGTTGATATT 300

Qy       1152 TATGGTGTAGAATATTACCATCTCTGTTATTTATGTGAAAATGCAGAAAGGAGTGCATGT 1211
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        301 TATGGTGTAGAATATTACCATCTCTGTTATTTATGTGAAAATGCAGAAAGGAGTGCATGT 360

Qy       1212 ATCTACGGAGGGGTAACAAATCATGAAGGGAATCATTTAGAAATTCCTAAAAAGATAGTC 1271
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        361 ATCTACGGAGGGGTAACAAATCATGAAGGGAATCATTTAGAAATTCCTAAAAAGATAGTC 420

Qy       1272 GTTAAAGTATCAATCGATGGTATCCAAAGCCTATCATTTGATATTGAAACAAATAAAAAA 1331
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        421 GTTAAAGTATCAATCGATGGTATCCAAAGCCTATCATTTGATATTGAAACAAATAAAAAA 480

Qy       1332 ATGGTAACTGCTCAAGAATTAGACTATAAAGTTAGAAAATATCTTACAGATAATAAGCAA 1391
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        481 ATGGTAACTGCTCAAGAATTAGACTATAAAGTTAGAAAATATCTTACAGATAATAAGCAA 540

Qy       1392 CTATATACTAATGGACCTTCTAAATATGAAACTGGATATATAAAGTTCATACCTAAGAAT 1451
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        541 CTATATACTAATGGACCTTCTAAATATGAAACTGGATATATAAAGTTCATACCTAAGAAT 600

Qy       1452 AAAGAAAGTTTTTGGTTTGATTTTTTCCCTGAACCAGAATTTACTCAATCTAAATATCTT 1511
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        601 AAAGAAAGTTTTTGGTTTGATTTTTTCCCTGAACCAGAATTTACTCAATCTAAATATCTT 660

Qy       1512 ATGATATATAAAGATAATGAAACGCTTGACTCAAACACAAGCCAAATT 1559
              ||||||||||||||||||||||||||||||||||||||||||||||||
Db        661 ATGATATATAAAGATAATGAAACGCTTGACTCAAACACAAGCCAAATT 708
```

RESULT 15
SP165SPEA

```
LOCUS       SP165SPEA                 708 bp    DNA     linear   BCT 23-OCT-1991
DEFINITION  S.pyogenes strain MGAS165 speA gene (allele 1) for type A exotoxin.
ACCESSION   X61556
VERSION     X61556.1  GI:47291
KEYWORDS    exotoxin; exotoxin type A.
SOURCE      Streptococcus pyogenes
  ORGANISM  Streptococcus pyogenes
            Bacteria; Firmicutes; Lactobacillales; Streptococcaceae;
            Streptococcus.
REFERENCE   1
  AUTHORS   Nelson,K., Schlievert,P.M., Selander,R.K. and Musser,J.M.
  TITLE     Characterization and clonal distribution of four alleles of the
            speA gene encoding pyrogenic exotoxin A (scarlet fever toxin) in
            Streptococcus pyogenes
  JOURNAL   J. Exp. Med. 174 (5), 1271-1274 (1991)
  MEDLINE   92044323
   PUBMED   1940804
REFERENCE   2  (bases 1 to 708)
  AUTHORS   Nelson,K.
  TITLE     Direct Submission
  JOURNAL   Submitted (01-SEP-1991) K. Nelson, Pennsylvania State University,
            Inst of Mol Evolutionary Genetics, University Park, PA 16802, USA
FEATURES             Location/Qualifiers
     source          1. .708
                     /organism="Streptococcus pyogenes"
                     /mol_type="genomic DNA"
                     /strain="MGAS165"
                     /isolate="Minnesota"
                     /db_xref="taxon:1314"
                     /chromosome="bacteriophage encoded"
                     /clone="direct PCR fragment"
     gene            1. .708
                     /gene="speA"
     CDS             <1. .>708
                     /gene="speA"
                     /codon_start=1
                     /transl_table=11
                     /product="type A exotoxin"
                     /protein_id="CAA43754.1"
                     /db_xref="GI:47292"
                     /db_xref="GOA:P97163"
                     /db_xref="SPTREMBL:P97163"
                     /translation="KKMVFFVLVTFLGLTISQEVFAQQDPDPSQLHRSSLVKNLQNIY
                     FLYEGDPVTHENVKSVDQLLSHDLIYNVSGPNYDKLKTELKNQEMATLFKDKNVDIYG
                     VEYYHLCYLCENAERSACIYGGVTNHEGNHLEIPKKIVVKVSIDGIQSLSFDIETNKK
                     MVTAQELDYKVRKYLTDNKQLYTNGPSKYETGYIKFIPKNKESFWFDFFPEPEFTQSK
                     YLMIYKDNETLDSNTSQI"
     sig_peptide     <1. .66
                     /gene="speA"
     mat_peptide     67. .>708
                     /gene="speA"
                     /product="type A exotoxin"
ORIGIN
```

```
  Query Match           38.2%;  Score 706.4;  DB 1;  Length 708;
  Best Local Similarity  99.9%;  Pred. No. 5e-92;
  Matches  707;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy        852 AAGAAAATGGTATTTTTTGTTTTAGTGACATTTCTTGGACTAACAATCTCGCAAGAGGTA 911
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 AAGAAAATGGTATTTTTTGTTTTAGTGACATTTCTTGGACTAACAATCTCGCAAGAGGTA 60

Qy        912 TTTGCTCAACAAGACCCCGATCCAAGCCAACTTCACAGATCTAGTTTAGTTAAAAACCTT 971
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 TTTGCTCAACAAGACCCCGATCCAAGCCAACTTCACAGATCTAGTTTAGTTAAAAACCTT 120

Qy        972 CAAAATATATATTTTCTTTATGAGGGTGACCCTGTTACTCACGAGAATGTGAAATCTGTT 1031
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 CAAAATATATATTTTCTTTATGAGGGTGACCCTGTTACTCACGAGAATGTGAAATCTGTT 180

Qy       1032 GATCAACTTTTATCTCACCATTTAATATATAATGTTTCAGGGCCAAATTATGATAAATTA 1091
              |||||||||||||||| ||||||||||||||||||||||||||||||||||||||||||
Db        181 GATCAACTTTTATCTCACGATTTAATATATAATGTTTCAGGGCCAAATTATGATAAATTA 240

Qy       1092 AAAACTGAACTTAAGAACCAAGAGATGGCAACTTTATTTAAGGATAAAAACGTTGATATT 1151
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 AAAACTGAACTTAAGAACCAAGAGATGGCAACTTTATTTAAGGATAAAAACGTTGATATT 300

Qy       1152 TATGGTGTAGAATATTACCATCTCTGTTATTTATGTGAAAATGCAGAAAGGAGTGCATGT 1211
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        301 TATGGTGTAGAATATTACCATCTCTGTTATTTATGTGAAAATGCAGAAAGGAGTGCATGT 360

Qy       1212 ATCTACGGAGGGGTAACAAATCATGAAGGGAATCATTTAGAAATTCCTAAAAAGATAGTC 1271
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        361 ATCTACGGAGGGGTAACAAATCATGAAGGGAATCATTTAGAAATTCCTAAAAAGATAGTC 420

Qy       1272 GTTAAAGTATCAATCGATGGTATCCAAAGCCTATCATTTGATATTGAAACAAATAAAAAA 1331
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        421 GTTAAAGTATCAATCGATGGTATCCAAAGCCTATCATTTGATATTGAAACAAATAAAAAA 480

Qy       1332 ATGGTAACTGCTCAAGAATTAGACTATAAAGTTAGAAAATATCTTACAGATAATAAGCAA 1391
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        481 ATGGTAACTGCTCAAGAATTAGACTATAAAGTTAGAAAATATCTTACAGATAATAAGCAA 540

Qy       1392 CTATATACTAATGGACCTTCTAAATATGAAACTGGATATATAAAGTTCATACCTAAGAAT 1451
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        541 CTATATACTAATGGACCTTCTAAATATGAAACTGGATATATAAAGTTCATACCTAAGAAT 600

Qy       1452 AAAGAAAGTTTTTGGTTTGATTTTTTCCCTGAACCAGAATTTACTCAATCTAAATATCTT 1511
```

```
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        601 AAAGAAAAGTTTTTGGTTTGATTTTTTTCCCTGAACCAGAATTTACTCAATCTAAATATCTT 660

Qy       1512 ATGATATATAAAGATAATGAAACGCTTGACTCAAACACAAGCCAAATT 1559
                ||||||||||||||||||||||||||||||||||||||||||||||||||
Db        661 ATGATATATAAAGATAATGAAACGCTTGACTCAAACACAAGCCAAATT 708
```

Search completed: May 9, 2004, 10:30:20
Job time : 7007 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: May 9, 2004, 04:48:51 ; Search time 701 Seconds
(without alignments)
11217.425 Million cell updates/sec

Title: US-10-625-221-12
Perfect score: 1851
Sequence: 1 ccatcacgcatcactcatgt.....ttagcaactattttatcgtc 1851

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 3373863 seqs, 2124099041 residues

Total number of hits satisfying chosen parameters: 6747726

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : N_Geneseq_29Jan04:*
1: geneseqn1980s:*
2: geneseqn1990s:*
3: geneseqn2000s:*
4: geneseqn2001as:*
5: geneseqn2001bs:*
6: geneseqn2002s:*
7: geneseqn2003as:*
8: geneseqn2003bs:*
9: geneseqn2003cs:*
10: geneseqn2004s:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

	Result No.	Score	% Query Match	Length	DB	ID	Description
	1	1851	100.0	1851	2	AAT51716	Aat51716 Streptoco
	2	1851	100.0	1851	2	AAV41593	Aav41593 Nucleotid
	3	1835.4	99.2	1837	2	AAT28540	Aat28540 S. pyogen
	4	1835.4	99.2	1837	4	ABA76857	Aba76857 Streptoco
	5	1832.2	99.0	1837	3	AAZ51112	Aaz51112 Streptoco
	6	1832.2	99.0	1837	6	ABN84229	Abn84229 Streptoco
	7	1832.2	99.0	1837	7	ACA61184	Aca61184 DNA encod
	8	1832.2	99.0	1837	8	AAD56771	Aad56771 Streptoco
	9	1830.6	98.9	1837	8	ACD28901	Acd28901 Streptoco
	10	892.8	48.2	1031	7	ACA64700	Aca64700 S. pyogen
	11	754.4	40.8	756	4	AAH01002	Aah01002 Unidentif
	12	656.8	35.5	1419	8	ACD28908	Acd28908 SPEa L42R
	13	656.8	35.5	1419	8	AAD56778	Aad56778 Streptoco
	14	278	15.0	1497	7	ACA64721	Aca64721 Staphyloc
	15	184.2	10.0	1712	3	AAZ51107	Aaz51107 Staphyloc
	16	184.2	10.0	1712	6	ABN84224	Abn84224 Staphyloc
	17	184.2	10.0	1712	7	ACA61179	Aca61179 DNA encod
	18	184.2	10.0	1712	7	ACA64695	Aca64695 S. aureus
	19	184.2	10.0	1712	8	ACD28896	Acd28896 S. aureus
	20	184.2	10.0	1712	8	AAD56766	Aad56766 DNA #3 re
c	21	178.8	9.7	110000	6	ABN71527_05	Continuation (6 of
	22	176.2	9.5	1712	3	AAZ51108	Aaz51108 Mutant St
	23	176.2	9.5	1712	6	ABN84225	Abn84225 Staphyloc
	24	176.2	9.5	1712	7	ACA61180	Aca61180 DNA encod
	25	176.2	9.5	1712	8	ACD28897	Acd28897 S. aureus
	26	176.2	9.5	1712	8	AAD56767	Aad56767 DNA #4 re
	27	171.4	9.3	773	2	AAT45698	Aat45698 Staphyloc
	28	171	9.2	886	7	ACA64689	Aca64689 S. aureus
	29	169.2	9.1	801	7	ACA64688	Aca64688 S. aureus
	30	167.8	9.1	1095	7	ACA64696	Aca64696 S. aureus
	31	164.6	8.9	1095	3	AAZ51111	Aaz51111 Staphyloc
	32	164.6	8.9	1095	6	ABN84228	Abn84228 Staphyloc
	33	164.6	8.9	1095	7	ACA61183	Aca61183 DNA encod
	34	164.6	8.9	1095	8	ACD28900	Acd28900 S. aureus
	35	164.6	8.9	1095	8	AAD56770	Aad56770 Staphyloc
	36	161.2	8.7	723	3	AAZ45835	Aaz45835 Mutant SE
	37	161.2	8.7	801	3	AAZ45834	Aaz45834 Mutant SE
	38	161.2	8.7	1388	3	AAZ51109	Aaz51109 Mutant St
	39	161.2	8.7	1388	6	ABN84226	Abn84226 Staphyloc
	40	161.2	8.7	1388	7	ACA61181	Aca61181 DNA encod
	41	161.2	8.7	1388	8	ACD28898	Acd28898 S. aureus
	42	161.2	8.7	1388	8	AAD56768	Aad56768 DNA #5 re
	43	159.6	8.6	867	4	AAH74983	Aah74983 Nucleotid
c	44	145.8	7.9	177	6	ABN66731	Abn66731 Streptoco
c	45	130.6	7.1	186	6	ABN66706	Abn66706 Streptoco

ALIGNMENTS

RESULT 1
AAT51716
ID AAT51716 standard; DNA; 1851 BP.
XX
AC AAT51716;
XX
DT 04-NOV-1997 (first entry)
XX
DE Streptococcus pyogenes Streptococcal toxin A DNA.
XX
KW Streptococcal; toxin A; SPE-A; non-lethal; mutant; production; vaccine;
KW protection; treatment; cancer; neutralising antibody;
KW streptococcal toxic shock syndrome; STSS; symptom; amelioration; fever;
KW hypotension; group A streptococcal infection; myositis; fascitis;
KW liver damage; T cell; lymphoma; ovarian; uterine; ss.
XX
OS Streptococcus pyogenes.
XX
FH Key Location/Qualifiers
FT CDS 828. .1583
FT /*tag= a
FT sig_peptide 828. .917
FT /*tag= b
FT mat_peptide 918. .1580
FT /*tag= c
FT /product= "Streptococcal_toxin_A"
XX
PN WO9640930-A1.
XX
PD 19-DEC-1996.
XX
PF 07-JUN-1996; 96WO-US010252.
XX
PR 07-JUN-1995; 95US-00480261.
XX
PA (MINU) UNIV MINNESOTA.
XX
PI Schlievert PM, Roggiani M, Stoehr J, Ohlendorf D;
XX
DR WPI; 1997-099936/09.
DR P-PSDB; AAW12097.
XX
PT Mutant SPE-A toxin with at least one amino acid change is substantially
PT non-lethal - used in vaccine composition for treatment of cancer and
PT streptococcal toxic shock syndrome etc.
XX

Applicant B

PS Disclosure; Page 77-79; 102pp; English.
XX
CC The present sequence encodes Streptococcus pyogenes Streptococcal toxin A
CC (SPE-A), from which a non-lethal mutant SPE-A, comprising at least 1
CC amino acid change, can be derived. The mutant SPE-A can be used to
CC produce vaccines to protect animals against wild type SPE-A and to treat
CC cancer and streptococcal toxic shock syndrome (STSS). The mutant SPE-A
CC causes neutralising antibodies (Ab) to be produced, which may be used to
CC ameliorate STSS symptoms, e.g. fever, hypotension, group A streptococcal
CC infection, myositis, fascitis and liver damage. The neutralising Ab is
CC preferably administered in conjunction with antibiotic therapy. The
CC mutant SPE-A is especially useful for treating T cell lymphomas, and
CC ovarian and uterine cancer. It is thought that mutant SPE-A can be
CC selectively toxic to T cell lymphoma cells
XX
SQ Sequence 1851 BP; 635 A; 326 C; 248 G; 642 T; 0 U; 0 Other;

Query Match 100.0%; Score 1851; DB 2; Length 1851;
Best Local Similarity 100.0%; Pred. No. 4.5e-289;
Matches 1851; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
Qy    1 CCATCACGCATCACTCATGTTTGACAGCTTATCATCGATAAGCTTACTTTTCGAATCAGG 60
Db    1 CCATCACGCATCACTCATGTTTGACAGCTTATCATCGATAAGCTTACTTTTCGAATCAGG 60

Qy   61 TCTATCCTTGAAACAGGTGCAACATAGATTAGGGCATGGAGATTTACCAGACAACTATGA 120
Db   61 TCTATCCTTGAAACAGGTGCAACATAGATTAGGGCATGGAGATTTACCAGACAACTATGA 120

Qy  121 ACGTATATACTCACATCACGCAATCGGCAATTGATGACATTGGAACTAAATTCAATCAAT 180
Db  121 ACGTATATACTCACATCACGCAATCGGCAATTGATGACATTGGAACTAAATTCAATCAAT 180

Qy  181 TTGTTACTAACAAGCAACTAGATTGACAACTAATTCTCAACAAACGTTAATTTAACAACA 240
Db  181 TTGTTACTAACAAGCAACTAGATTGACAACTAATTCTCAACAAACGTTAATTTAACAACA 240

Qy  241 TTCAAGTAACTCCCACCAGCTCCATCAATGCTTACCGTAAGTAATCATAACTTACTAAAA 300
Db  241 TTCAAGTAACTCCCACCAGCTCCATCAATGCTTACCGTAAGTAATCATAACTTACTAAAA 300

Qy  301 CCTTGTTACATCAAGGTTTTTTCTTTTTGTCTTGTTCATGAGTTACCATAACTTTCTATA 360
Db  301 CCTTGTTACATCAAGGTTTTTTCTTTTTGTCTTGTTCATGAGTTACCATAACTTTCTATA 360

Qy  361 TTATTGACAACTAAATTGACAACTCTTCAATTATTTTTCTGTCTACTCAAAGTTTTCTTC 420
Db  361 TTATTGACAACTAAATTGACAACTCTTCAATTATTTTTCTGTCTACTCAAAGTTTTCTTC 420

Qy  421 ATTTGATATAGTCTAATTCCACCATCACTTCTTCCACTCTCTCTACCGTCACAACTTCAT 480
Db  421 ATTTGATATAGTCTAATTCCACCATCACTTCTTCCACTCTCTCTACCGTCACAACTTCAT 480

Qy  481 CATCTCTCACTTTTTCGTGTGGTAACACATAATCAAATATCTTTCCGTTTTTACGCACTA 540
Db  481 CATCTCTCACTTTTTCGTGTGGTAACACATAATCAAATATCTTTCCGTTTTTACGCACTA 540

Qy  541 TCGCTACTGTGTCACCTAAAATATACCCCTTATCAATCGCTTCTTTAAACTCATCTATAT 600
Db  541 TCGCTACTGTGTCACCTAAAATATACCCCTTATCAATCGCTTCTTTAAACTCATCTATAT 600

Qy  601 ATAACATATTTCATCCTCCTACCTATCTATTCGTAAAAAGATAAAAATAACTATTGTTTT 660
Db  601 ATAACATATTTCATCCTCCTACCTATCTATTCGTAAAAAGATAAAAATAACTATTGTTTT 660

Qy  661 TTTTGTTATTTTATAATAAAATTATTAATATAAGTTAATGTTTTTTAAAAATATACAATT 720
Db  661 TTTTGTTATTTTATAATAAAATTATTAATATAAGTTAATGTTTTTTAAAAATATACAATT 720

Qy  721 TTATTCTATTTATAGTTAGCTATTTTTTCATTGTTAGTAATATTGGTGAATTGTAATAAC 780
Db  721 TTATTCTATTTATAGTTAGCTATTTTTTCATTGTTAGTAATATTGGTGAATTGTAATAAC 780

Qy  781 CTTTTTAAATCTAGAGGAGAACCCAGATATAAAATGGAGGAATATTAATGGAAAACAATA 840
Db  781 CTTTTTAAATCTAGAGGAGAACCCAGATATAAAATGGAGGAATATTAATGGAAAACAATA 840

Qy  841 AAAAAGTATTGAAGAAAATGGTATTTTTTGTTTTAGTGACATTTCTTGGACTAACAATCT 900
Db  841 AAAAAGTATTGAAGAAAATGGTATTTTTTGTTTTAGTGACATTTCTTGGACTAACAATCT 900

Qy  901 CGCAAGAGGTATTTGCTCAACAAGACCCCGATCCAAGCCAACTTCACAGATCTAGTTTAG 960
Db  901 CGCAAGAGGTATTTGCTCAACAAGACCCCGATCCAAGCCAACTTCACAGATCTAGTTTAG 960

Qy  961 TTAAAAACCTTCAAAATATATATTTTCTTTATGAGGGTGACCCTGTTACTCACGAGAATG 1020
Db  961 TTAAAAACCTTCAAAATATATATTTTCTTTATGAGGGTGACCCTGTTACTCACGAGAATG 1020

Qy 1021 TGAAATCTGTTGATCAACTTTTATCTCACCATTTAATATATAATGTTTCAGGGCCAAATT 1080
Db 1021 TGAAATCTGTTGATCAACTTTTATCTCACCATTTAATATATAATGTTTCAGGGCCAAATT 1080

Qy 1081 ATGATAAATTAAAAACTGAACTTAAGAACCAAGAGATGGCAACTTTATTTAAGGATAAAA 1140
Db 1081 ATGATAAATTAAAAACTGAACTTAAGAACCAAGAGATGGCAACTTTATTTAAGGATAAAA 1140

Qy 1141 ACGTTGATATTTATGGTGTAGAATATTACCATCTCTGTTATTTATGTGAAAATGCAGAAA 1200
Db 1141 ACGTTGATATTTATGGTGTAGAATATTACCATCTCTGTTATTTATGTGAAAATGCAGAAA 1200

Qy 1201 GGAGTGCATGTATCTACGGAGGGGTAACAAATCATGAAGGGAATCATTTAGAAATTCCTA 1260
Db 1201 GGAGTGCATGTATCTACGGAGGGGTAACAAATCATGAAGGGAATCATTTAGAAATTCCTA 1260

Qy 1261 AAAAGATAGTCGTTAAAGTATCAATCGATGGTATCCAAAGCCTATCATTTGATATTGAAA 1320
Db 1261 AAAAGATAGTCGTTAAAGTATCAATCGATGGTATCCAAAGCCTATCATTTGATATTGAAA 1320

Qy 1321 CAAATAAAAAAAATGGTAACTGCTCAAGAATTAGACTATAAAGTTAGAAAATATCTTACAG 1380
Db 1321 CAAATAAAAAAAATGGTAACTGCTCAAGAATTAGACTATAAAGTTAGAAAATATCTTACAG 1380

Qy 1381 ATAATAAGCAACTATATACTAATGGACCTTCTAAATATGAAACTGGATATATAAAGTTCA 1440
Db 1381 ATAATAAGCAACTATATACTAATGGACCTTCTAAATATGAAACTGGATATATAAAGTTCA 1440

Qy 1441 TACCTAAGAATAAAGAAAGTTTTTGGTTTGATTTTTTCCCTGAACCAGAATTTACTCAAT 1500
Db 1441 TACCTAAGAATAAAGAAAGTTTTTGGTTTGATTTTTTCCCTGAACCAGAATTTACTCAAT 1500

Qy 1501 CTAAATATCTTATGATATATAAAGATAATGAAACGCTTGACTCAAACACAAGCCAAATTG 1560
Db 1501 CTAAATATCTTATGATATATAAAGATAATGAAACGCTTGACTCAAACACAAGCCAAATTG 1560

Qy 1561 AAGTCTACCTAACAACCAAGTAACTTTTTGCTTTTGGCAACCTTACCTACTGCTGGATTT 1620
Db 1561 AAGTCTACCTAACAACCAAGTAACTTTTTGCTTTTGGCAACCTTACCTACTGCTGGATTT 1620

Qy 1621 AGAAATTTTATTGCAATTCTTTTATTAATGTAAAAACCGCTCATTTGATGAGCGGTTTTG 1680
Db 1621 AGAAATTTTATTGCAATTCTTTTATTAATGTAAAAACCGCTCATTTGATGAGCGGTTTTG 1680

Qy 1681 TCTTATCTAAAGGAGCTTTACCTCCTAATGCTGCAAAATTTTAAATGTTGGATTTTTGTA 1740
Db 1681 TCTTATCTAAAGGAGCTTTACCTCCTAATGCTGCAAAATTTTAAATGTTGGATTTTTGTA 1740

Qy 1741 TTTGTCTATTGTATTTGATGGGTAATCCCATTTTTCGACAGACATCGTCGTGCCACCTCT 1800
Db 1741 TTTGTCTATTGTATTTGATGGGTAATCCCATTTTTCGACAGACATCGTCGTGCCACCTCT 1800

Qy 1801 AACACCAAAATCATAGACAGGAGCTTGTAGCTTAGCAACTATTTTATCGTC 1851
Db 1801 AACACCAAAATCATAGACAGGAGCTTGTAGCTTAGCAACTATTTTATCGTC 1851
```

RESULT 2
AAV41593
ID AAV41593 standard; DNA; 1851 BP.
XX
AC AAV41593;
XX
DT 12-OCT-1998 (first entry)
XX
DE Nucleotide sequence of Streptococcus pyogenes exotoxin A.
XX
KW SPE-A toxin; nonlethal; mutant; Streptococcus pyogenes exotoxin A;
KW wild type; nontoxic; antibody; vaccine; immunity; ovarian cancer;
KW streptococcal toxic shock syndrome; STSS; T cell lymphoma;
KW uterine cancer; ss.
XX
OS Streptococcus pyogenes.
XX
FH Key Location/Qualifiers
FT CDS 828. .1583
FT /*tag= a
FT /product= "SPE-A toxin"
XX
PN WO9824911-A2.
XX
PD 11-JUN-1998.
XX
PF 05-DEC-1997; 97WO-US022228.
XX
PR 06-DEC-1996; 96US-0032930P.
XX
PA (MINU) UNIV MINNESOTA.
XX
PI Schlievert PM, Roggiani M, Stoehr J, Ohlendorf D;
XX
DR WPI; 1998-333330/29.
DR P-PSDB; AAW59780.
XX
PT New mutant Streptococcal SPE-A toxins - useful for, e.g. prevention or
PT treatment of streptococcal infection or toxic shock syndrome.
XX
PS Disclosure; Fig 3; 95pp; English.
XX
CC This is the nucleotide sequence of the Streptococcus pyogenes exotoxin A
CC (SPE-A toxin). The novel mutant Streptococcal SPE-A toxin has at least 1
CC aa change and is nonlethal compared with a protein to wild type SPE-A
CC toxin. The mutant SPE-A toxins are nontoxic and can produce antibodies
CC that neutralise wild type SPE-A toxin activity. The toxins can be used in
CC vaccines and therapeutics to generate a protective immune response
CC against streptococcal infection. They can be used to protect against the
CC development of streptococcal toxic shock syndrome (STSS). In addition,
CC the toxins can be used for treating animals with symptoms of
CC streptococcal infection or STSS and in methods for stimulating T cell
CC proliferation and in the treatment of cancer. In particular they can be
CC used for treating T cell lymphomas, and ovarian and uterine cancer
XX
SQ Sequence 1851 BP; 635 A; 326 C; 248 G; 642 T; 0 U; 0 Other;

Query Match 100.0%; Score 1851; DB 2; Length 1851;
Best Local Similarity 100.0%; Pred. No. 4.5e-289;
Matches 1851; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
Qy    1 CCATCACGCATCACTCATGTTTGACAGCTTATCATCGATAAGCTTACTTTTCGAATCAGG 60
Db    1 CCATCACGCATCACTCATGTTTGACAGCTTATCATCGATAAGCTTACTTTTCGAATCAGG 60

Qy   61 TCTATCCTTGAAACAGGTGCAACATAGATTAGGGCATGGAGATTTACCAGACAACTATGA 120
Db   61 TCTATCCTTGAAACAGGTGCAACATAGATTAGGGCATGGAGATTTACCAGACAACTATGA 120

Qy  121 ACGTATATACTCACATCACGCAATCGGCAATTGATGACATTGGAACTAAATTCAATCAAT 180
Db  121 ACGTATATACTCACATCACGCAATCGGCAATTGATGACATTGGAACTAAATTCAATCAAT 180

Qy  181 TTGTTACTAACAAGCAACTAGATTGACAACTAATTCTCAACAAACGTTAATTTAACAACA 240
Db  181 TTGTTACTAACAAGCAACTAGATTGACAACTAATTCTCAACAAACGTTAATTTAACAACA 240

Qy  241 TTCAAGTAACTCCCACCAGCTCCATCAATGCTTACCGTAAGTAATCATAACTTACTAAAA 300
Db  241 TTCAAGTAACTCCCACCAGCTCCATCAATGCTTACCGTAAGTAATCATAACTTACTAAAA 300

Qy  301 CCTTGTTACATCAAGGTTTTTTCTTTTTGTCTTGTTCATGAGTTACCATAACTTTCTATA 360
Db  301 CCTTGTTACATCAAGGTTTTTTCTTTTTGTCTTGTTCATGAGTTACCATAACTTTCTATA 360

Qy  361 TTATTGACAACTAAATTGACAACTCTTCAATTATTTTTCTGTCTACTCAAAGTTTTCTTC 420
Db  361 TTATTGACAACTAAATTGACAACTCTTCAATTATTTTTCTGTCTACTCAAAGTTTTCTTC 420

Qy  421 ATTTGATATAGTCTAATTCCACCATCACTTCTTCCACTCTCTCTACCGTCACAACTTCAT 480
Db  421 ATTTGATATAGTCTAATTCCACCATCACTTCTTCCACTCTCTCTACCGTCACAACTTCAT 480

Qy  481 CATCTCTCACTTTTTCGTGTGGTAACACATAATCAAATATCTTTCCGTTTTTACGCACTA 540
Db  481 CATCTCTCACTTTTTCGTGTGGTAACACATAATCAAATATCTTTCCGTTTTTACGCACTA 540

Qy  541 TCGCTACTGTGTCACCTAAAATATACCCCTTATCAATCGCTTCTTTAAACTCATCTATAT 600
Db  541 TCGCTACTGTGTCACCTAAAATATACCCCTTATCAATCGCTTCTTTAAACTCATCTATAT 600

Qy  601 ATAACATATTTCATCCTCCTACCTATCTATTCGTAAAAAGATAAAAATAACTATTGTTTT 660
Db  601 ATAACATATTTCATCCTCCTACCTATCTATTCGTAAAAAGATAAAAATAACTATTGTTTT 660

Qy  661 TTTTGTTATTTTATAATAAAATTATTAATATAAGTTAATGTTTTTTAAAAATATACAATT 720
Db  661 TTTTGTTATTTTATAATAAAATTATTAATATAAGTTAATGTTTTTTAAAAATATACAATT 720

Qy  721 TTATTCTATTTATAGTTAGCTATTTTTTCATTGTTAGTAATATTGGTGAATTGTAATAAC 780
Db  721 TTATTCTATTTATAGTTAGCTATTTTTTCATTGTTAGTAATATTGGTGAATTGTAATAAC 780

Qy  781 CTTTTTAAATCTAGAGGAGAACCCAGATATAAAATGGAGGAATATTAATGGAAAACAATA 840
Db  781 CTTTTTAAATCTAGAGGAGAACCCAGATATAAAATGGAGGAATATTAATGGAAAACAATA 840

Qy  841 AAAAAGTATTGAAGAAAATGGTATTTTTTGTTTTAGTGACATTTCTTGGACTAACAATCT 900
Db  841 AAAAAGTATTGAAGAAAATGGTATTTTTTGTTTTAGTGACATTTCTTGGACTAACAATCT 900

Qy  901 CGCAAGAGGTATTTGCTCAACAAGACCCCGATCCAAGCCAACTTCACAGATCTAGTTTAG 960
Db  901 CGCAAGAGGTATTTGCTCAACAAGACCCCGATCCAAGCCAACTTCACAGATCTAGTTTAG 960

Qy  961 TTAAAAACCTTCAAAATATATATTTTCTTTATGAGGGTGACCCTGTTACTCACGAGAATG 1020
Db  961 TTAAAAACCTTCAAAATATATATTTTCTTTATGAGGGTGACCCTGTTACTCACGAGAATG 1020

Qy 1021 TGAAATCTGTTGATCAACTTTTATCTCACCATTTAATATATAATGTTTCAGGGCCAAATT 1080
Db 1021 TGAAATCTGTTGATCAACTTTTATCTCACCATTTAATATATAATGTTTCAGGGCCAAATT 1080

Qy 1081 ATGATAAATTAAAAACTGAACTTAAGAACCAAGAGATGGCAACTTTATTTAAGGATAAAA 1140
Db 1081 ATGATAAATTAAAAACTGAACTTAAGAACCAAGAGATGGCAACTTTATTTAAGGATAAAA 1140

Qy 1141 ACGTTGATATTTATGGTGTAGAATATTACCATCTCTGTTATTTATGTGAAAATGCAGAAA 1200
Db 1141 ACGTTGATATTTATGGTGTAGAATATTACCATCTCTGTTATTTATGTGAAAATGCAGAAA 1200

Qy 1201 GGAGTGCATGTATCTACGGAGGGGTAACAAATCATGAAGGGAATCATTTAGAAATTCCTA 1260
Db 1201 GGAGTGCATGTATCTACGGAGGGGTAACAAATCATGAAGGGAATCATTTAGAAATTCCTA 1260

Qy 1261 AAAAGATAGTCGTTAAAGTATCAATCGATGGTATCCAAAGCCTATCATTTGATATTGAAA 1320
```

```
Db       1261 AAAAGATAGTCGTTAAAGTATCAATCGATGGTATCCAAAGCCTATCATTTGATATTGAAA 1320

Qy       1321 CAAATAAAAAAATGGTAACTGCTCAAGAATTAGACTATAAAGTTAGAAAATATCTTACAG 1380
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1321 CAAATAAAAAAATGGTAACTGCTCAAGAATTAGACTATAAAGTTAGAAAATATCTTACAG 1380

Qy       1381 ATAATAAGCAACTATATACTAATGGACCTTCTAAATATGAAACTGGATATATAAAGTTCA 1440
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1381 ATAATAAGCAACTATATACTAATGGACCTTCTAAATATGAAACTGGATATATAAAGTTCA 1440

Qy       1441 TACCTAAGAATAAAGAAAGTTTTTGGTTTGATTTTTTCCCTGAACCAGAATTTACTCAAT 1500
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1441 TACCTAAGAATAAAGAAAGTTTTTGGTTTGATTTTTTCCCTGAACCAGAATTTACTCAAT 1500

Qy       1501 CTAAATATCTTATGATATATAAAGATAATGAAACGCTTGACTCAAACACAAGCCAAATTG 1560
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1501 CTAAATATCTTATGATATATAAAGATAATGAAACGCTTGACTCAAACACAAGCCAAATTG 1560

Qy       1561 AAGTCTACCTAACAACCAAGTAACTTTTTGCTTTTGGCAACCTTACCTACTGCTGGATTT 1620
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1561 AAGTCTACCTAACAACCAAGTAACTTTTTGCTTTTGGCAACCTTACCTACTGCTGGATTT 1620

Qy       1621 AGAAATTTTATTGCAATTCTTTTATTAATGTAAAAACCGCTCATTTGATGAGCGGTTTTG 1680
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1621 AGAAATTTTATTGCAATTCTTTTATTAATGTAAAAACCGCTCATTTGATGAGCGGTTTTG 1680

Qy       1681 TCTTATCTAAAGGAGCTTTACCTCCTAATGCTGCAAAATTTTAAATGTTGGATTTTTGTA 1740
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1681 TCTTATCTAAAGGAGCTTTACCTCCTAATGCTGCAAAATTTTAAATGTTGGATTTTTGTA 1740

Qy       1741 TTTGTCTATTGTATTTGATGGGTAATCCCATTTTTCGACAGACATCGTCGTGCCACCTCT 1800
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1741 TTTGTCTATTGTATTTGATGGGTAATCCCATTTTTCGACAGACATCGTCGTGCCACCTCT 1800

Qy       1801 AACACCAAAATCATAGACAGGAGCTTGTAGCTTAGCAACTATTTTATCGTC 1851
             |||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1801 AACACCAAAATCATAGACAGGAGCTTGTAGCTTAGCAACTATTTTATCGTC 1851
```

RESULT 3
AAT28540
ID AAT28540 standard; DNA; 1837 BP.
XX
AC AAT28540;
XX
DT 02-APR-1997 (first entry)
XX
DE S. pyogenes detection probe #2.
XX
KW Detection; probe; amplification primer; bacterial pathogen; pneumonia;
KW Escherichia coli; Klebsiella pneumoniae; Pseudomonas aeruginosa;
KW Proteus mirabilis; Streptococcus pneumoniae; Staphylococcus aureus;
KW Staphylococcus epidermidis; Enterococcus faecalis; respiratory tract;
KW Staphylococcus saprophyticus; Streptococcus pyogenes; urinary tract;
KW Haemophilus influenzae; Moraxella catarrhalis; septicaemia; meningitis;
KW infection; intra-abdominal infection; skin infection;
KW bacterial resistance; beta-lactam antibiotic; ds.
XX
OS Synthetic.
XX
PN WO9608582-A2.
XX
PD 21-MAR-1996.
XX
PF 12-SEP-1995; 95WO-CA000528.
XX
PR 12-SEP-1994; 94US-00304732.
XX
PA (BERG/) BERGERON M G.
PA (OUEL/) OUELLETTE M.
PA (ROYP/) ROY P H.
XX
PI Bergeron MG, Ouellette M, Roy PH;
XX
DR WPI; 1996-179953/18.
XX
PT Method for the detection of bacterial species using probes and primers -
PT allows detection and quantification of antibiotic resistant bacteria in
PT patients, the environment and food.
XX
PS Claim 57; Page 92-93; 216pp; English.
XX
CC The sequences given in AAT28539-40 represent probes which were used in
CC the method of the invention for the detection of S. pyogenes in a sample.
CC The method comprises using probes and/or amplification primers which are
CC specific, ubiquitous and sensitive for determining the presence and/or
CC amount of nucleic acids from selected bacterial species in any sample,
CC where the bacterial nucleic acid comprises a selected target region
CC hybridisable with the probes or primers. The method comprises contacting
CC the sample with the probes or primers and detecting the presence and/or
CC amount of hybridised primers or amplification products as and indication
CC of the presence and/or amount of the bacterial species. This method may
CC be used to detect commonly encountered bacterial pathogens, e.g.
CC Escherichia coli, Klebsiella pneumoniae, Pseudomonas aeruginosa, Proteus
CC mirabilis, Streptococcus pneumoniae, Staphylococcus aureus,
CC Staphylococcus epidermidis, Enterococcus faecalis, Staphylococcus
CC saprophyticus, Streptococcus pyogenes, Haemophilus influenzae and
CC Moraxella catarrhalis. These bacterial species are associated with
CC approx. 90% of urinary tract infections and with a high percentage of
CC other severe infections including septicaemia, meningitis, pneumonia,
CC intra-abdominal infections, skin infections and other severe respiratory
CC tract infections. The method may also be used to evaluate a bacterial
CC resistance to beta-lactam antibiotics
XX
SQ Sequence 1837 BP; 631 A; 318 C; 248 G; 640 T; 0 U; 0 Other;

Query Match 99.2%; Score 1835.4; DB 2; Length 1837;
Best Local Similarity 99.9%; Pred. No. 1.4e-286;
Matches 1836; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

```
Qy         15 TCATGTTTGACAGCTTATCATCGATAAGCTTACTTTTCGAATCAGGTCTATCCTTGAAAC 74
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 TCATGTTTGACAGCTTATCATCGATAAGCTTACTTTTCGAATCAGGTCTATCCTTGAAAC 60

Qy         75 AGGTGCAACATAGATTAGGGCATGGAGATTTACCAGACAACTATGAACGTATATACTCAC 134
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 AGGTGCAACATAGATTAGGGCATGGAGATTTACCAGACAACTATGAACGTATATACTCAC 120

Qy        135 ATCACGCAATCGGCAATTGATGACATTGGAACTAAATTCAATCAATTTGTTACTAACAAG 194
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 ATCACGCAATCGGCAATTGATGACATTGGAACTAAATTCAATCAATTTGTTACTAACAAG 180

Qy        195 CAACTAGATTGACAACTAATTCTCAACAAACGTTAATTTAACAACATTCAAGTAACTCCC 254
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 CAACTAGATTGACAACTAATTCTCAACAAACGTTAATTTAACAACATTCAAGTAACTCCC 240

Qy        255 ACCAGCTCCATCAATGCTTACCGTAAGTAATCATAACTTACTAAAACCTTGTTACATCAA 314
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 ACCAGCTCCATCAATGCTTACCGTAAGTAATCATAACTTACTAAAACCTTGTTACATCAA 300

Qy        315 GGTTTTTTCTTTTTGTCTTGTTCATGAGTTACCATAACTTTCTATATTATTGACAACTAA 374
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        301 GGTTTTTTCTTTTTGTCTTGTTCATGAGTTACCATAACTTTCTATATTATTGACAACTAA 360

Qy        375 ATTGACAACTCTTCAATTATTTTTCTGTCTACTCAAAGTTTTCTTCATTTGATATAGTCT 434
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        361 ATTGACAACTCTTCAATTATTTTTCTGTCTACTCAAAGTTTTCTTCATTTGATATAGTCT 420

Qy        435 AATTCCACCATCACTTCTTCCACTCTCTCTACCGTCACAACTTCATCATCTCTCACTTTT 494
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        421 AATTCCACCATCACTTCTTCCACTCTCTCTACCGTCACAACTTCATCATCTCTCACTTTT 480

Qy        495 TCGTGTGGTAACACATAATCAAATATCTTTCCGTTTTTACGCACTATCGCTACTGTGTCA 554
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        481 TCGTGTGGTAACACATAATCAAATATCTTTCCGTTTTTACGCACTATCGCTACTGTGTCA 540
```

```
Qy       555 CCTAAAATATACCCCTTATCAATCGCTTCTTTAAACTCATCTATATATAACATATTTCAT 614
Db       541 CCTAAAATATACCCCTTATCAATCGCTTCTTTAAACTCATCTATATATAACATATTTCAT 600

Qy       615 CCTCCTACCTATCTATTCGTAAAAAGATAAAAATAACTATTGTTTTTTTTGTTATTTTAT 674
Db       601 CCTCCTACCTATCTATTCGTAAAAAGATAAAAATAACTATTGTTTTTTTTGTTATTTTAT 660

Qy       675 AATAAAATTATTAATATAAGTTAATGTTTTTTAAAAATATACAATTTTATTCTATTTATA 734
Db       661 AATAAAATTATTAATATAAGTTAATGTTTTTTAAAAATATACAATTTTATTCTATTTATA 720

Qy       735 GTTAGCTATTTTTTCATTGTTAGTAATATTGGTGAATTGTAATAACCTTTTTAAATCTAG 794
Db       721 GTTAGCTATTTTTTCATTGTTAGTAATATTGGTGAATTGTAATAACCTTTTTAAATCTAG 780

Qy       795 AGGAGAACCCAGATATAAAATGGAGGAATATTAATGGAAAACAATAAAAAAGTATTGAAG 854
Db       781 AGGAGAACCCAGATATAAAATGGAGGAATATTAATGGAAAACAATAAAAAAGTATTGAAG 840

Qy       855 AAAATGGTATTTTTTGTTTTAGTGACATTTCTTGGACTAACAATCTCGCAAGAGGTATTT 914
Db       841 AAAATGGTATTTTTTGTTTTAGTGACATTTCTTGGACTAACAATCTCGCAAGAGGTATTT 900

Qy       915 GCTCAACAAGACCCCGATCCAAGCCAACTTCACAGATCTAGTTTAGTTAAAAACCTTCAA 974
Db       901 GCTCAACAAGACCCCGATCCAAGCCAACTTCACAGATCTAGTTTAGTTAAAAACCTTCAA 960

Qy       975 AATATATATTTTCTTTATGAGGGTGACCCTGTTACTCACGAGAATGTGAAATCTGTTGAT 1034
Db       961 AATATATATTTTCTTTATGAGGGTGACCCTGTTACTCACGAGAATGTGAAATCTGTTGAT 1020

Qy      1035 CAACTTTTATCTCACCATTTAATATATAATGTTTCAGGGCCAAATTATGATAAATTAAAA 1094
Db      1021 CAACTTTTATCTCACGATTTAATATATAATGTTTCAGGGCCAAATTATGATAAATTAAAA 1080

Qy      1095 ACTGAACTTAAGAACCAAGAGATGGCAACTTTATTTAAGGATAAAAACGTTGATATTTAT 1154
Db      1081 ACTGAACTTAAGAACCAAGAGATGGCAACTTTATTTAAGGATAAAAACGTTGATATTTAT 1140

Qy      1155 GGTGTAGAATATTACCATCTCTGTTATTTATGTGAAAATGCAGAAAGGAGTGCATGTATC 1214
Db      1141 GGTGTAGAATATTACCATCTCTGTTATTTATGTGAAAATGCAGAAAGGAGTGCATGTATC 1200

Qy      1215 TACGGAGGGGTAACAAATCATGAAGGGAATCATTTAGAAATTCCTAAAAAGATAGTCGTT 1274
Db      1201 TACGGAGGGGTAACAAATCATGAAGGGAATCATTTAGAAATTCCTAAAAAGATAGTCGTT 1260

Qy      1275 AAAGTATCAATCGATGGTATCCAAAGCCTATCATTTGATATTGAAACAAATAAAAAAATG 1334
Db      1261 AAAGTATCAATCGATGGTATCCAAAGCCTATCATTTGATATTGAAACAAATAAAAAAATG 1320

Qy      1335 GTAACTGCTCAAGAATTAGACTATAAAGTTAGAAAATATCTTACAGATAATAAGCAACTA 1394
Db      1321 GTAACTGCTCAAGAATTAGACTATAAAGTTAGAAAATATCTTACAGATAATAAGCAACTA 1380

Qy      1395 TATACTAATGGACCTTCTAAATATGAAACTGGATATATAAAGTTCATACCTAAGAATAAA 1454
Db      1381 TATACTAATGGACCTTCTAAATATGAAACTGGATATATAAAGTTCATACCTAAGAATAAA 1440

Qy      1455 GAAAGTTTTTGGTTTGATTTTTTCCCTGAACCAGAATTTACTCAATCTAAATATCTTATG 1514
Db      1441 GAAAGTTTTTGGTTTGATTTTTTCCCTGAACCAGAATTTACTCAATCTAAATATCTTATG 1500

Qy      1515 ATATATAAAGATAATGAAACGCTTGACTCAAACACAAGCCAAATTGAAGTCTACCTAACA 1574
Db      1501 ATATATAAAGATAATGAAACGCTTGACTCAAACACAAGCCAAATTGAAGTCTACCTAACA 1560

Qy      1575 ACCAAGTAACTTTTTGCTTTTGGCAACCTTACCTACTGCTGGATTTAGAAATTTTATTGC 1634
Db      1561 ACCAAGTAACTTTTTGCTTTTGGCAACCTTACCTACTGCTGGATTTAGAAATTTTATTGC 1620

Qy      1635 AATTCTTTTATTAATGTAAAAACCGCTCATTTGATGAGCGGTTTTGTCTTATCTAAAGGA 1694
Db      1621 AATTCTTTTATTAATGTAAAAACCGCTCATTTGATGAGCGGTTTTGTCTTATCTAAAGGA 1680

Qy      1695 GCTTTACCTCCTAATGCTGCAAAATTTTAAATGTTGGATTTTTGTATTTGTCTATTGTAT 1754
Db      1681 GCTTTACCTCCTAATGCTGCAAAATTTTAAATGTTGGATTTTTGTATTTGTCTATTGTAT 1740

Qy      1755 TTGATGGGTAATCCCATTTTTCGACAGACATCGTCGTGCCACCTCTAACACCAAAATCAT 1814
Db      1741 TTGATGGGTAATCCCATTTTTCGACAGACATCGTCGTGCCACCTCTAACACCAAAATCAT 1800

Qy      1815 AGACAGGAGCTTGTAGCTTAGCAACTATTTTATCGTC 1851
Db      1801 AGACAGGAGCTTGTAGCTTAGCAACTATTTTATCGTC 1837
```

RESULT 4
ABA76857
ID ABA76857 standard; DNA; 1837 BP.
XX
AC ABA76857;
XX
DT 28-JAN-2002 (first entry)
XX
DE Streptococcus pyogenes polynucleotide SEQ ID NO 33.
XX
KW Detection; bacterial species; animal; food; environment;
KW antibiotic resistance; ds.
XX
OS Streptococcus pyogenes.
XX
PN NZ501596-A.
XX
PD 29-JUN-2001.
XX
PF 12-SEP-1995; 95NZ-00501596.
XX
PR 12-SEP-1995; 95NZ-00501596.
XX
PA (IDII-) IDI INFECTIO DIAGNOSTIC INC.
XX
PI Bergeron MG, Ouellette M, Roy PH;
XX
DR WPI; 2001-615034/71.
XX
PT Method for detecting target bacterial species in a sample, comprises
PT detecting the presence or amount of bacterial nucleic acid amplified by a
PT primer derived from bacterial DNA, specific for the target bacterial
PT species.
XX
PS Claim 6; Page 107-108; 168pp; English.
XX
CC The invention relates to detecting target bacterial species suspected to
CC be present in a sample, comprising contacting nucleic acids of target
CC bacterial species with an amplification primer pair derived from a
CC bacterial DNA fragment (ABA76825-ABA76861) specific for the target
CC bacterial species but ubiquitous for different strains, amplifying the
CC nucleic acid and detecting the presence or amount of an amplified
CC sequence as an indication of the presence or amount of the target
CC bacterial species. The invention includes primers and probes (ABA76862-
CC ABA76984) against the target bacterial species, espeially E.coli,
CC K.pneumoniae, P.aeruginosa, P.mirabilis, S.pneumoniae, S.aureus,
CC S.epidermidis, E.faecalis, S.saprophyticus, S.pyogenes, H.influenzae,
CC M.catarrhalis and/or group A Streptococci producing exotoxin A gene spe
CC A, suspected to be present in a sample which is obtained from human
CC patients, animals, environment or food, and which consists of one or more
CC bacterial colonies. Oligonucleotide probes and primers complementary to
CC the bacterial genes encoding resistance to antibiotics such as bla(tem),
CC bla(rob), bla(shv), aadB, aacC1, aacC2, aacC3, aacA4, mecA, vanA, vanH,
CC vanX, satA, aacA-aphD, vat, vga, msrA, sul and/or int (ABA76985-ABA77001)
CC are also useful to identify commonly encountered and clinically important
CC resistance genes. The invention provides a rapid method of bacterial

```
CC   identification that can be achieved, which reduces the time currently
CC   required for the identification of pathogens in the clinical laboratory
XX
SQ   Sequence 1837 BP; 631 A; 318 C; 248 G; 640 T; 0 U; 0 Other;

  Query Match             99.2%;  Score 1835.4;  DB 4;  Length 1837;
  Best Local Similarity   99.9%;  Pred. No. 1.4e-286;
  Matches 1836;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy         15 TCATGTTTGACAGCTTATCATCGATAAGCTTACTTTTCGAATCAGGTCTATCCTTGAAAC  74
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 TCATGTTTGACAGCTTATCATCGATAAGCTTACTTTTCGAATCAGGTCTATCCTTGAAAC  60

Qy         75 AGGTGCAACATAGATTAGGGCATGGAGATTTACCAGACAACTATGAACGTATATACTCAC  134
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 AGGTGCAACATAGATTAGGGCATGGAGATTTACCAGACAACTATGAACGTATATACTCAC  120

Qy        135 ATCACGCAATCGGCAATTGATGACATTGGAACTAAATTCAATCAATTTGTTACTAACAAG  194
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 ATCACGCAATCGGCAATTGATGACATTGGAACTAAATTCAATCAATTTGTTACTAACAAG  180

Qy        195 CAACTAGATTGACAACTAATTCTCAACAAACGTTAATTTAACAACATTCAAGTAACTCCC  254
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 CAACTAGATTGACAACTAATTCTCAACAAACGTTAATTTAACAACATTCAAGTAACTCCC  240

Qy        255 ACCAGCTCCATCAATGCTTACCGTAAGTAATCATAACTTACTAAAACCTTGTTACATCAA  314
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 ACCAGCTCCATCAATGCTTACCGTAAGTAATCATAACTTACTAAAACCTTGTTACATCAA  300

Qy        315 GGTTTTTTCTTTTTGTCTTGTTCATGAGTTACCATAACTTTCTATATTATTGACAACTAA  374
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        301 GGTTTTTTCTTTTTGTCTTGTTCATGAGTTACCATAACTTTCTATATTATTGACAACTAA  360

Qy        375 ATTGACAACTCTTCAATTATTTTTCTGTCTACTCAAAGTTTTCTTCATTTGATATAGTCT  434
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        361 ATTGACAACTCTTCAATTATTTTTCTGTCTACTCAAAGTTTTCTTCATTTGATATAGTCT  420

Qy        435 AATTCCACCATCACTTCTTCCACTCTCTCTACCGTCACAACTTCATCATCTCTCACTTTT  494
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        421 AATTCCACCATCACTTCTTCCACTCTCTCTACCGTCACAACTTCATCATCTCTCACTTTT  480

Qy        495 TCGTGTGGTAACACATAATCAAATATCTTTCCGTTTTTACGCACTATCGCTACTGTGTCA  554
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        481 TCGTGTGGTAACACATAATCAAATATCTTTCCGTTTTTACGCACTATCGCTACTGTGTCA  540

Qy        555 CCTAAAATATACCCCTTATCAATCGCTTCTTTAAACTCATCTATATATAACATATTTCAT  614
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        541 CCTAAAATATACCCCTTATCAATCGCTTCTTTAAACTCATCTATATATAACATATTTCAT  600

Qy        615 CCTCCTACCTATCTATTCGTAAAAAGATAAAAATAACTATTGTTTTTTTTGTTATTTTAT  674
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        601 CCTCCTACCTATCTATTCGTAAAAAGATAAAAATAACTATTGTTTTTTTTGTTATTTTAT  660

Qy        675 AATAAAATTATTAATATAAGTTAATGTTTTTTAAAAATATACAATTTTATTCTATTTATA  734
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        661 AATAAAATTATTAATATAAGTTAATGTTTTTTAAAAATATACAATTTTATTCTATTTATA  720

Qy        735 GTTAGCTATTTTTTTCATTGTTAGTAATATTGGTGAATTGTAATAACCTTTTTAAATCTAG  794
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        721 GTTAGCTATTTTTTTCATTGTTAGTAATATTGGTGAATTGTAATAACCTTTTTAAATCTAG  780

Qy        795 AGGAGAACCCAGATATAAAATGGAGGAATATTAATGGAAAACAATAAAAAAGTATTGAAG  854
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        781 AGGAGAACCCAGATATAAAATGGAGGAATATTAATGGAAAACAATAAAAAAGTATTGAAG  840

Qy        855 AAAATGGTATTTTTTGTTTTAGTGACATTTCTTGGACTAACAATCTCGCAAGAGGTATTT  914
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        841 AAAATGGTATTTTTTGTTTTAGTGACATTTCTTGGACTAACAATCTCGCAAGAGGTATTT  900

Qy        915 GCTCAACAAGACCCCGATCCAAGCCAACTTCACAGATCTAGTTTAGTTAAAAACCTTCAA  974
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        901 GCTCAACAAGACCCCGATCCAAGCCAACTTCACAGATCTAGTTTAGTTAAAAACCTTCAA  960

Qy        975 AATATATATTTTCTTTATGAGGGTGACCCTGTTACTCACGAGAATGTGAAATCTGTTGAT 1034
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        961 AATATATATTTTCTTTATGAGGGTGACCCTGTTACTCACGAGAATGTGAAATCTGTTGAT 1020

Qy       1035 CAACTTTTATCTCACCATTTAATATATAATGTTTCAGGGCCAAATTATGATAAATTAAAA 1094
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1021 CAACTTTTATCTCACGATTTAATATATAATGTTTCAGGGCCAAATTATGATAAATTAAAA 1080

Qy       1095 ACTGAACTTAAGAACCAAGAGATGGCAACTTTATTTAAGGATAAAAACGTTGATATTTAT 1154
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1081 ACTGAACTTAAGAACCAAGAGATGGCAACTTTATTTAAGGATAAAAACGTTGATATTTAT 1140

Qy       1155 GGTGTAGAATATTACCATCTCTGTTATTTATGTGAAAATGCAGAAAGGAGTGCATGTATC 1214
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1141 GGTGTAGAATATTACCATCTCTGTTATTTATGTGAAAATGCAGAAAGGAGTGCATGTATC 1200

Qy       1215 TACGGAGGGGTAACAAATCATGAAGGGAATCATTTAGAAATTCCTAAAAAGATAGTCGTT 1274
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1201 TACGGAGGGGTAACAAATCATGAAGGGAATCATTTAGAAATTCCTAAAAAGATAGTCGTT 1260

Qy       1275 AAAGTATCAATCGATGGTATCCAAAGCCTATCATTTGATATTGAAACAAATAAAAAAATG 1334
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1261 AAAGTATCAATCGATGGTATCCAAAGCCTATCATTTGATATTGAAACAAATAAAAAAATG 1320

Qy       1335 GTAACTGCTCAAGAATTAGACTATAAAGTTAGAAAATATCTTACAGATAATAAGCAACTA 1394
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1321 GTAACTGCTCAAGAATTAGACTATAAAGTTAGAAAATATCTTACAGATAATAAGCAACTA 1380

Qy       1395 TATACTAATGGACCTTCTAAATATGAAACTGGATATATAAAGTTCATACCTAAGAATAAA 1454
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1381 TATACTAATGGACCTTCTAAATATGAAACTGGATATATAAAGTTCATACCTAAGAATAAA 1440

Qy       1455 GAAAGTTTTTGGTTTGATTTTTTCCCTGAACCAGAATTTACTCAATCTAAATATCTTATG 1514
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1441 GAAAGTTTTTGGTTTGATTTTTTCCCTGAACCAGAATTTACTCAATCTAAATATCTTATG 1500

Qy       1515 ATATATAAAGATAATGAAACGCTTGACTCAAACACAAGCCAAATTGAAGTCTACCTAACA 1574
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1501 ATATATAAAGATAATGAAACGCTTGACTCAAACACAAGCCAAATTGAAGTCTACCTAACA 1560

Qy       1575 ACCAAGTAACTTTTTGCTTTTGGCAACCTTACCTACTGCTGGATTTAGAAATTTTATTGC 1634
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1561 ACCAAGTAACTTTTTGCTTTTGGCAACCTTACCTACTGCTGGATTTAGAAATTTTATTGC 1620

Qy       1635 AATTCTTTTATTAATGTAAAAACCGCTCATTTGATGAGCGGTTTTGTCTTATCTAAAGGA 1694
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1621 AATTCTTTTATTAATGTAAAAACCGCTCATTTGATGAGCGGTTTTGTCTTATCTAAAGGA 1680

Qy       1695 GCTTTACCTCCTAATGCTGCAAAATTTTAAATGTTGGATTTTTGTATTTGTCTATTGTAT 1754
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1681 GCTTTACCTCCTAATGCTGCAAAATTTTAAATGTTGGATTTTTGTATTTGTCTATTGTAT 1740

Qy       1755 TTGATGGGTAATCCCATTTTTCGACAGACATCGTCGTGCCACCTCTAACACCAAAATCAT 1814
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1741 TTGATGGGTAATCCCATTTTTCGACAGACATCGTCGTGCCACCTCTAACACCAAAATCAT 1800

Qy       1815 AGACAGGAGCTTGTAGCTTAGCAACTATTTTATCGTC 1851
               |||||||||||||||||||||||||||||||||||||
Db       1801 AGACAGGAGCTTGTAGCTTAGCAACTATTTTATCGTC 1837

RESULT 5
AAZ51112
ID   AAZ51112 standard; DNA; 1837 BP.
XX
AC   AAZ51112;
XX
DT   05-JUN-2000  (first entry)
XX
DE   Streptococcal pyrogenic exotoxin A (SPE-A) encoding DNA.
XX
KW   Superantigen toxin; SAg; Streptococcal pyrogenic exotoxin A; SPE-A;
KW   antibacterial; vaccine; MHC class II receptor; T-cell antigen receptor;
```

```
KW   cytostatic; antibody; staphylococcal/streptococcal toxin; toxoid; SPEA42;
KW   diagnosis; treatment; superantigen-associated bacterial infection; ds.
XX
OS   Streptococcus sp.
XX
FH   Key             Location/Qualifiers
FT   CDS             814. .1569
FT                   /*tag= a
FT                   /product= "Streptococcal pyrogenic exotoxin A (SPE-A)"
XX
PN   WO200009154-A1.
XX
PD   24-FEB-2000.
XX
PF   13-AUG-1998;   98WO-US016766.
XX
PR   13-AUG-1998;   98WO-US016766.
XX
PA   (REED-) REED ARMY INST RES WALTER.
XX
PI   Ulrich RG,  Olson MA,  Bavari S;
XX
DR   WPI; 2000-224177/19.
DR   P-PSDB; AAY70109.
XX
PT   Nucleic acid encoding superantigen toxin useful as a vaccine and for
PT   diagnosis of superantigen-associated bacterial infections.
XX
PS   Example 12; Page 92-94; 118pp; English.
XX
CC   The present sequence is the DNA encoding Streptococcal pyrogenic exotoxin
CC   A (SPE-A), a bacterial superantigen toxin (SAg), used for the formulation
CC   of SPE-A vaccine SPEA42. The coding region of this SAg toxin is altered
CC   by site directed mutagenesis, introducing L42R mutation, that results in
CC   disruption of binding of the toxin to both the MHC class II or T-cell
CC   antigen receptor. SPE-A has antibacterial and cytostatic activity. This
CC   sequence is useful for the production of SPE-A vaccines and specific
CC   antibodies. This vaccine overcomes the disadvantages of the chemically
CC   inactivated toxoids and is designed to protect individuals against one or
CC   several related staphylococcal and streptococcal toxins. It is used for
CC   the diagnosis and treatment or amelioration of superantigen-associated
CC   bacterial infections
XX
SQ   Sequence 1837 BP; 632 A; 318 C; 249 G; 638 T; 0 U; 0 Other;

  Query Match             99.0%;  Score 1832.2;  DB 3;  Length 1837;
  Best Local Similarity   99.8%;  Pred. No. 4.7e-286;
  Matches 1834;  Conservative    0;  Mismatches    3;  Indels    0;  Gaps    0;

Qy         15 TCATGTTTGACAGCTTATCATCGATAAGCTTACTTTTCGAATCAGGTCTATCCTTGAAAC 74
Db          1 TCATGTTTGACAGCTTATCATCGATAAGCTTACTTTTCGAATCAGGTCTATCCTTGAAAC 60

Qy         75 AGGTGCAACATAGATTAGGGCATGGAGATTTACCAGACAACTATGAACGTATATACTCAC 134
Db         61 AGGTGCAACATAGATTAGGGCATGGAGATTTACCAGACAACTATGAACGTATATACTCAC 120

Qy        135 ATCACGCAATCGGCAATTGATGACATTGGAACTAAATTCAATCAATTTGTTACTAACAAG 194
Db        121 ATCACGCAATCGGCAATTGATGACATTGGAACTAAATTCAATCAATTTGTTACTAACAAG 180

Qy        195 CAACTAGATTGACAACTAATTCTCAACAAACGTTAATTTAACAACATTCAAGTAACTCCC 254
Db        181 CAACTAGATTGACAACTAATTCTCAACAAACGTTAATTTAACAACATTCAAGTAACTCCC 240

Qy        255 ACCAGCTCCATCAATGCTTACCGTAAGTAATCATAACTTACTAAAACCTTGTTACATCAA 314
Db        241 ACCAGCTCCATCAATGCTTACCGTAAGTAATCATAACTTACTAAAACCTTGTTACATCAA 300

Qy        315 GGTTTTTTCTTTTTGTCTTGTTCATGAGTTACCATAACTTTCTATATTATTGACAACTAA 374
Db        301 GGTTTTTTCTTTTTGTCTTGTTCATGAGTTACCATAACTTTCTATATTATTGACAACTAA 360

Qy        375 ATTGACAACTCTTCAATTATTTTTCTGTCTACTCAAAGTTTTCTTCATTTGATATAGTCT 434
Db        361 ATTGACAACTCTTCAATTATTTTTCTGTCTACTCAAAGTTTTCTTCATTTGATATAGTCT 420

Qy        435 AATTCCACCATCACTTCTTCCACTCTCTCTACCGTCACAACTTCATCATCTCTCACTTTT 494
Db        421 AATTCCACCATCACTTCTTCCACTCTCTCTACCGTCACAACTTCATCATCTCTCACTTTT 480

Qy        495 TCGTGTGGTAACACATAATCAAATATCTTTCCGTTTTTACGCACTATCGCTACTGTGTCA 554
Db        481 TCGTGTGGTAACACATAATCAAATATCTTTCCGTTTTTACGCACTATCGCTACTGTGTCA 540

Qy        555 CCTAAAATATACCCCTTATCAATCGCTTCTTTAAACTCATCTATATATAACATATTTCAT 614
Db        541 CCTAAAATATACCCCTTATCAATCGCTTCTTTAAACTCATCTATATATAACATATTTCAT 600

Qy        615 CCTCCTACCTATCTATTCGTAAAAAGATAAAAATAACTATTGTTTTTTTTGTTATTTTAT 674
Db        601 CCTCCTACCTATCTATTCGTAAAAAGATAAAAATAACTATTGTTTTTTTTGTTATTTTAT 660

Qy        675 AATAAAATTATTAATATAAGTTAATGTTTTTTAAAAATATACAATTTTATTCTATTTATA 734
Db        661 AATAAAATTATTAATATAAGTTAATGTTTTTTAAAAATATACAATTTTATTCTATTTATA 720

Qy        735 GTTAGCTATTTTTTCATTGTTAGTAATATTGGTGAATTGTAATAACCTTTTTAAATCTAG 794
Db        721 GTTAGCTATTTTTTCATTGTTAGTAATATTGGTGAATTGTAATAACCTTTTTAAATCTAG 780

Qy        795 AGGAGAACCCAGATATAAAATGGAGGAATATTAATGGAAAACAATAAAAAAGTATTGAAG 854
Db        781 AGGAGAACCCAGATATAAAATGGAGGAATATTAATGGAAAACAATAAAAAAGTATTGAAG 840

Qy        855 AAAATGGTATTTTTTGTTTTAGTGACATTTCTTGGACTAACAATCTCGCAAGAGGTATTT 914
Db        841 AAAATGGTATTTTTTGTTTTAGTGACATTTCTTGGACTAACAATCTCGCAAGAGGTATTT 900

Qy        915 GCTCAACAAGACCCCGATCCAAGCCAACTTCACAGATCTAGTTTAGTTAAAAACCTTCAA 974
Db        901 GCTCAACAAGACCCCGATCCAAGCCAACTTCACAGATCTAGTTTAGTTAAAAACCTTCAA 960

Qy        975 AATATATATTTTCTTTATGAGGGTGACCCTGTTACTCACGAGAATGTGAAATCTGTTGAT 1034
Db        961 AATATATATTTTCTTTATGAGGGTGACCCTGTTACTCACGAGAATGTGAAATCTGTTGAT 1020

Qy       1035 CAACTTTTATCTCACCATTTAATATATAATGTTTCAGGGCCAAATTATGATAAATTAAAA 1094
Db       1021 CAACTTAGATCTCACGATTTAATATATAATGTTTCAGGGCCAAATTATGATAAATTAAAA 1080

Qy       1095 ACTGAACTTAAGAACCAAGAGATGGCAACTTTATTTAAGGATAAAAACGTTGATATTTAT 1154
Db       1081 ACTGAACTTAAGAACCAAGAGATGGCAACTTTATTTAAGGATAAAAACGTTGATATTTAT 1140

Qy       1155 GGTGTAGAATATTACCATCTCTGTTATTTATGTGAAAATGCAGAAAGGAGTGCATGTATC 1214
Db       1141 GGTGTAGAATATTACCATCTCTGTTATTTATGTGAAAATGCAGAAAGGAGTGCATGTATC 1200

Qy       1215 TACGGAGGGGTAACAAATCATGAAGGGAATCATTTAGAAATTCCTAAAAAGATAGTCGTT 1274
Db       1201 TACGGAGGGGTAACAAATCATGAAGGGAATCATTTAGAAATTCCTAAAAAGATAGTCGTT 1260

Qy       1275 AAAGTATCAATCGATGGTATCCAAAGCCTATCATTTGATATTGAAACAAATAAAAAAATG 1334
Db       1261 AAAGTATCAATCGATGGTATCCAAAGCCTATCATTTGATATTGAAACAAATAAAAAAATG 1320

Qy       1335 GTAACTGCTCAAGAATTAGACTATAAAGTTAGAAAATATCTTACAGATAATAAGCAACTA 1394
Db       1321 GTAACTGCTCAAGAATTAGACTATAAAGTTAGAAAATATCTTACAGATAATAAGCAACTA 1380

Qy       1395 TATACTAATGGACCTTCTAAATATGAAACTGGATATATAAAGTTCATACCTAAGAATAAA 1454
Db       1381 TATACTAATGGACCTTCTAAATATGAAACTGGATATATAAAGTTCATACCTAAGAATAAA 1440

Qy       1455 GAAAGTTTTTGGTTTGATTTTTTCCCTGAACCAGAATTTACTCAATCTAAATATCTTATG 1514
```

```
Db       1441 GAAAGTTTTTGGTTTGATTTTTTCCCTGAACCAGAATTTACTCAATCTAAATATCTTATG 1500

Qy       1515 ATATATAAAGATAATGAAACGCTTGACTCAAACACAAGCCAAATTGAAGTCTACCTAACA 1574

Db       1501 ATATATAAAGATAATGAAACGCTTGACTCAAACACAAGCCAAATTGAAGTCTACCTAACA 1560

Qy       1575 ACCAAGTAACTTTTTGCTTTTGGCAACCTTACCTACTGCTGGATTTAGAAATTTTATTGC 1634

Db       1561 ACCAAGTAACTTTTTGCTTTTGGCAACCTTACCTACTGCTGGATTTAGAAATTTTATTGC 1620

Qy       1635 AATTCTTTTATTAATGTAAAAACCGCTCATTTGATGAGCGGTTTTGTCTTATCTAAAGGA 1694

Db       1621 AATTCTTTTATTAATGTAAAAACCGCTCATTTGATGAGCGGTTTTGTCTTATCTAAAGGA 1680

Qy       1695 GCTTTACCTCCTAATGCTGCAAAATTTTAAATGTTGGATTTTTGTATTTGTCTATTGTAT 1754

Db       1681 GCTTTACCTCCTAATGCTGCAAAATTTTAAATGTTGGATTTTTGTATTTGTCTATTGTAT 1740

Qy       1755 TTGATGGGTAATCCCATTTTTCGACAGACATCGTCGTGCCACCTCTAACACCAAAATCAT 1814

Db       1741 TTGATGGGTAATCCCATTTTTCGACAGACATCGTCGTGCCACCTCTAACACCAAAATCAT 1800

Qy       1815 AGACAGGAGCTTGTAGCTTAGCAACTATTTTATCGTC 1851

Db       1801 AGACAGGAGCTTGTAGCTTAGCAACTATTTTATCGTC 1837
```

RESULT 6
ABN84229
ID ABN84229 standard; cDNA; 1837 BP.
XX
AC ABN84229;
XX
DT 23-SEP-2002 (first entry)
XX
DE Streptococcal pyrogenic exotoxin A vaccine SPEa42 coding sequence.
XX
KW Exotoxin A; SPEa; SPEa42; superantigen; antigen; toxin; vaccine;
KW attenuation; mutant; gene; ss.
XX
OS Streptococcus pyogenes.
OS Synthetic.
XX
FH Key Location/Qualifiers
FT CDS 814. .1569
FT /*tag= a
FT /product= "SPEa42"
XX
PN US6399332-B1.
XX
PD 04-JUN-2002.
XX
PF 01-SEP-1998; 98US-00144776.
XX
PR 25-JUN-1997; 97US-00882431.
XX
PA (USSA) US SEC OF ARMY.
XX
PI Ulrich RG, Olson MA, Bavari S;
XX
DR WPI; 2002-546281/58.
DR P-PSDB; ABB79508.
XX
PT Novel isolated and purified superantigen toxin DNA fragment which has
PT been genetically altered, useful for producing vaccine for treatment of
PT superantigen toxin-associated bacterial diseases.
XX
PS Disclosure; Col 61-63; 46pp; English.
XX
CC The present sequence is the coding sequence of Streptococcus pyogenes
CC vaccine SPEa42. The vaccine differs from the native SPEa sequence by
CC substitution of the Leu-42 residue by Arg. This mutation is expected to
CC disrupt contact between the toxin and the HLA-DR receptor, reducing DR1
CC binding. SPEa42 can be expressed as a recombinant protein in Escherichia
CC coli as a secreted protein or as a cytoplasmic product. No indicators of
CC toxicity have been detected for the purified recombinant protein, and
CC vaccine studies demonstrate that SPEa42 is highly antigenic, inducing
CC protective immunity in a mouse animal model. The attenuated superantigen
CC can be used to protect against superantigen toxin infections. Methods of
CC producing and using altered superantigen toxins as vaccines, and in
CC diagnosis and therapy, are provided by the invention. A multivalent
CC vaccine consisting of altered superantigen toxins from SEA, SEB, SEC-1,
CC TSST-1 and SPEa is predicted to provide protective immunity against the
CC majority of bacterial superantigen toxins
XX
SQ Sequence 1837 BP; 632 A; 318 C; 249 G; 638 T; 0 U; 0 Other;

Query Match 99.0%; Score 1832.2; DB 6; Length 1837;
Best Local Similarity 99.8%; Pred. No. 4.7e-286;
Matches 1834; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

```
Qy         15 TCATGTTTGACAGCTTATCATCGATAAGCTTACTTTTCGAATCAGGTCTATCCTTGAAAC 74

Db          1 TCATGTTTGACAGCTTATCATCGATAAGCTTACTTTTCGAATCAGGTCTATCCTTGAAAC 60

Qy         75 AGGTGCAACATAGATTAGGGCATGGAGATTTACCAGACAACTATGAACGTATATACTCAC 134

Db         61 AGGTGCAACATAGATTAGGGCATGGAGATTTACCAGACAACTATGAACGTATATACTCAC 120

Qy        135 ATCACGCAATCGGCAATTGATGACATTGGAACTAAATTCAATCAATTTGTTACTAACAAG 194

Db        121 ATCACGCAATCGGCAATTGATGACATTGGAACTAAATTCAATCAATTTGTTACTAACAAG 180

Qy        195 CAACTAGATTGACAACTAATTCTCAACAAACGTTAATTTAACAACATTCAAGTAACTCCC 254

Db        181 CAACTAGATTGACAACTAATTCTCAACAAACGTTAATTTAACAACATTCAAGTAACTCCC 240

Qy        255 ACCAGCTCCATCAATGCTTACCGTAAGTAATCATAACTTACTAAAACCTTGTTACATCAA 314

Db        241 ACCAGCTCCATCAATGCTTACCGTAAGTAATCATAACTTACTAAAACCTTGTTACATCAA 300

Qy        315 GGTTTTTTCTTTTTGTCTTGTTCATGAGTTACCATAACTTTCTATATTATTGACAACTAA 374

Db        301 GGTTTTTTCTTTTTGTCTTGTTCATGAGTTACCATAACTTTCTATATTATTGACAACTAA 360

Qy        375 ATTGACAACTCTTCAATTATTTTTCTGTCTACTCAAAGTTTTCTTCATTTGATATAGTCT 434

Db        361 ATTGACAACTCTTCAATTATTTTTCTGTCTACTCAAAGTTTTCTTCATTTGATATAGTCT 420

Qy        435 AATTCCACCATCACTTCTTCCACTCTCTCTACCGTCACAACTTCATCATCTCTCACTTTT 494

Db        421 AATTCCACCATCACTTCTTCCACTCTCTCTACCGTCACAACTTCATCATCTCTCACTTTT 480

Qy        495 TCGTGTGGTAACACATAATCAAATATCTTTCCGTTTTTACGCACTATCGCTACTGTGTCA 554

Db        481 TCGTGTGGTAACACATAATCAAATATCTTTCCGTTTTTACGCACTATCGCTACTGTGTCA 540

Qy        555 CCTAAAATATACCCCTTATCAATCGCTTCTTTAAACTCATCTATATATAACATATTTCAT 614

Db        541 CCTAAAATATACCCCTTATCAATCGCTTCTTTAAACTCATCTATATATAACATATTTCAT 600

Qy        615 CCTCCTACCTATCTATTCGTAAAAAGATAAAAATAACTATTGTTTTTTTTGTTATTTTAT 674

Db        601 CCTCCTACCTATCTATTCGTAAAAAGATAAAAATAACTATTGTTTTTTTTGTTATTTTAT 660

Qy        675 AATAAAATTATTAATATAAGTTAATGTTTTTTAAAAATATACAATTTTATTCTATTTATA 734

Db        661 AATAAAATTATTAATATAAGTTAATGTTTTTTAAAAATATACAATTTTATTCTATTTATA 720

Qy        735 GTTAGCTATTTTTTCATTGTTAGTAATATTGGTGAATTGTAATAACCTTTTTAAATCTAG 794

Db        721 GTTAGCTATTTTTTCATTGTTAGTAATATTGGTGAATTGTAATAACCTTTTTAAATCTAG 780

Qy        795 AGGAGAACCCAGATATAAAATGGAGGAATATTAATGGAAAACAATAAAAAAGTATTGAAG 854
```

```
Db        781 AGGAGAACCCAGATATAAAATGGAGGAATATTAATGGAAAACAATAAAAAAGTATTGAAG 840
Qy        855 AAAATGGTATTTTTTGTTTTAGTGACATTTCTTGGACTAACAATCTCGCAAGAGGTATTT 914
Db        841 AAAATGGTATTTTTTGTTTTAGTGACATTTCTTGGACTAACAATCTCGCAAGAGGTATTT 900
Qy        915 GCTCAACAAGACCCCGATCCAAGCCAACTTCACAGATCTAGTTTAGTTAAAAACCTTCAA 974
Db        901 GCTCAACAAGACCCCGATCCAAGCCAACTTCACAGATCTAGTTTAGTTAAAAACCTTCAA 960
Qy        975 AATATATATTTTCTTTATGAGGGTGACCCTGTTACTCACGAGAATGTGAAATCTGTTGAT 1034
Db        961 AATATATATTTTCTTTATGAGGGTGACCCTGTTACTCACGAGAATGTGAAATCTGTTGAT 1020
Qy       1035 CAACTTTTATCTCACCATTTAATATATAATGTTTCAGGGCCAAATTATGATAAATTAAAA 1094
Db       1021 CAACTTAGATCTCACGATTTAATATATAATGTTTCAGGGCCAAATTATGATAAATTAAAA 1080
Qy       1095 ACTGAACTTAAGAACCAAGAGATGGCAACTTTATTTAAGGATAAAAACGTTGATATTTAT 1154
Db       1081 ACTGAACTTAAGAACCAAGAGATGGCAACTTTATTTAAGGATAAAAACGTTGATATTTAT 1140
Qy       1155 GGTGTAGAATATTACCATCTCTGTTATTTATGTGAAAATGCAGAAAGGAGTGCATGTATC 1214
Db       1141 GGTGTAGAATATTACCATCTCTGTTATTTATGTGAAAATGCAGAAAGGAGTGCATGTATC 1200
Qy       1215 TACGGAGGGGTAACAAATCATGAAGGGAATCATTTAGAAATTCCTAAAAAGATAGTCGTT 1274
Db       1201 TACGGAGGGGTAACAAATCATGAAGGGAATCATTTAGAAATTCCTAAAAAGATAGTCGTT 1260
Qy       1275 AAAGTATCAATCGATGGTATCCAAAGCCTATCATTTGATATTGAAACAAATAAAAAAATG 1334
Db       1261 AAAGTATCAATCGATGGTATCCAAAGCCTATCATTTGATATTGAAACAAATAAAAAAATG 1320
Qy       1335 GTAACTGCTCAAGAATTAGACTATAAAGTTAGAAAATATCTTACAGATAATAAGCAACTA 1394
Db       1321 GTAACTGCTCAAGAATTAGACTATAAAGTTAGAAAATATCTTACAGATAATAAGCAACTA 1380
Qy       1395 TATACTAATGGACCTTCTAAATATGAAACTGGATATATAAAGTTCATACCTAAGAATAAA 1454
Db       1381 TATACTAATGGACCTTCTAAATATGAAACTGGATATATAAAGTTCATACCTAAGAATAAA 1440
Qy       1455 GAAAGTTTTTGGTTTGATTTTTTCCCTGAACCAGAATTTACTCAATCTAAATATCTTATG 1514
Db       1441 GAAAGTTTTTGGTTTGATTTTTTCCCTGAACCAGAATTTACTCAATCTAAATATCTTATG 1500
Qy       1515 ATATATAAAGATAATGAAACGCTTGACTCAAACACAAGCCAAATTGAAGTCTACCTAACA 1574
Db       1501 ATATATAAAGATAATGAAACGCTTGACTCAAACACAAGCCAAATTGAAGTCTACCTAACA 1560
Qy       1575 ACCAAGTAACTTTTTGCTTTTGGCAACCTTACCTACTGCTGGATTTAGAAATTTTATTGC 1634
Db       1561 ACCAAGTAACTTTTTGCTTTTGGCAACCTTACCTACTGCTGGATTTAGAAATTTTATTGC 1620
Qy       1635 AATTCTTTTATTAATGTAAAAACCGCTCATTTGATGAGCGGTTTTGTCTTATCTAAAGGA 1694
Db       1621 AATTCTTTTATTAATGTAAAAACCGCTCATTTGATGAGCGGTTTTGTCTTATCTAAAGGA 1680
Qy       1695 GCTTTACCTCCTAATGCTGCAAAATTTTAAATGTTGGATTTTTGTATTTGTCTATTGTAT 1754
Db       1681 GCTTTACCTCCTAATGCTGCAAAATTTTAAATGTTGGATTTTTGTATTTGTCTATTGTAT 1740
Qy       1755 TTGATGGGTAATCCCATTTTTCGACAGACATCGTCGTGCCACCTCTAACACCAAAATCAT 1814
Db       1741 TTGATGGGTAATCCCATTTTTCGACAGACATCGTCGTGCCACCTCTAACACCAAAATCAT 1800
Qy       1815 AGACAGGAGCTTGTAGCTTAGCAACTATTTTATCGTC 1851
Db       1801 AGACAGGAGCTTGTAGCTTAGCAACTATTTTATCGTC 1837
```

RESULT 7
ACA61184

```
ID   ACA61184 standard; DNA; 1837 BP.
XX
AC   ACA61184;
XX
DT   11-AUG-2003  (first entry)
XX
DE   DNA encoding streptococcal pyrogenic exotoxin A.
XX
KW   Pyrogenic exotoxin A; ds; superantigen-associated bacterial infection;
KW   superantigen toxin; gene; vaccine.
XX
OS   Streptococcus sp.
XX
FH   Key             Location/Qualifiers
FT   CDS             814...1569
FT                   /*tag= a
FT                   /product= "Pyrogenic exotoxin A"
XX
PN   US2003009015-A1.
XX
PD   09-JAN-2003.
XX
PF   25-JUN-1997;   97US-00882431.
XX
PR   25-JUN-1997;   97US-00882431.
XX
PA   (ULRI/) ULRICH R G.
PA   (OLSO/) OLSON M A.
PA   (BAVA/) BAVARI S.
XX
PI   Ulrich RG, Olson MA, Bavari S;
XX
DR   WPI; 2003-401542/38.
DR   P-PSDB; ABU10088.
XX
PT   New superantigen toxin and/or DNA fragment with an altered binding of the
PT   encoded altered toxin to either MHC class II or T cell antigen receptor,
PT   useful for treating or ameliorating superantigen-associated bacterial
PT   infection.
XX
PS   Claim 9; Page 32-34; 50pp; English.
XX
CC   The invention relates to an isolated and purified superantigen toxin
CC   and/or DNA fragment, which has been altered so that the binding of the
CC   encoded toxin to either major histocompatibility complex (MHC) class II
CC   or T cell antigen receptor is altered. The superantigen toxins, DNA
CC   fragments, and vaccines are useful for treating or ameliorating
CC   superantigen-associated bacterial infection. The DNA fragments are
CC   particularly useful for producing vaccine against superantigen toxin
CC   infections. The transformed host cells are useful for analysing the
CC   effectiveness of drugs and agents that affect the binding of
CC   superantigens to MHC class II or T-cell antigen receptors. The present
CC   sequence represents DNA encoding streptococcal pyrogenic exotoxin A
XX
SQ   Sequence 1837 BP; 632 A; 318 C; 249 G; 638 T; 0 U; 0 Other;
```

Query Match 99.0%; Score 1832.2; DB 7; Length 1837;
Best Local Similarity 99.8%; Pred. No. 4.7e-286;
Matches 1834; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

```
Qy        15 TCATGTTTGACAGCTTATCATCGATAAGCTTACTTTTCGAATCAGGTCTATCCTTGAAAC 74
Db         1 TCATGTTTGACAGCTTATCATCGATAAGCTTACTTTTCGAATCAGGTCTATCCTTGAAAC 60
Qy        75 AGGTGCAACATAGATTAGGGCATGGAGATTTACCAGACAACTATGAACGTATATACTCAC 134
Db        61 AGGTGCAACATAGATTAGGGCATGGAGATTTACCAGACAACTATGAACGTATATACTCAC 120
Qy       135 ATCACGCAATCGGCAATTGATGACATTGGAACTAAATTCAATCAATTTGTTACTAACAAG 194
Db       121 ATCACGCAATCGGCAATTGATGACATTGGAACTAAATTCAATCAATTTGTTACTAACAAG 180
Qy       195 CAACTAGATTGACAACTAATTCTCAACAAACGTTAATTTAACAACATTCAAGTAACTCCC 254
```

```
Db       181 CAACTAGATTGACAACTAATTCTCAACAAACGTTAATTTAACAACATTCAAGTAACTCCC 240
Qy       255 ACCAGCTCCATCAATGCTTACCGTAAGTAATCATAACTTACTAAAACCTTGTTACATCAA 314
Db       241 ACCAGCTCCATCAATGCTTACCGTAAGTAATCATAACTTACTAAAACCTTGTTACATCAA 300
Qy       315 GGTTTTTTCTTTTTGTCTTGTTCATGAGTTACCATAACTTTCTATATTATTGACAACTAA 374
Db       301 GGTTTTTTCTTTTTGTCTTGTTCATGAGTTACCATAACTTTCTATATTATTGACAACTAA 360
Qy       375 ATTGACAACTCTTCAATTATTTTTCTGTCTACTCAAAGTTTTCTTCATTTGATATAGTCT 434
Db       361 ATTGACAACTCTTCAATTATTTTTCTGTCTACTCAAAGTTTTCTTCATTTGATATAGTCT 420
Qy       435 AATTCCACCATCACTTCTTCCACTCTCTCTACCGTCACAACTTCATCATCTCTCACTTTT 494
Db       421 AATTCCACCATCACTTCTTCCACTCTCTCTACCGTCACAACTTCATCATCTCTCACTTTT 480
Qy       495 TCGTGTGGTAACACATAATCAAATATCTTTCCGTTTTTACGCACTATCGCTACTGTGTCA 554
Db       481 TCGTGTGGTAACACATAATCAAATATCTTTCCGTTTTTACGCACTATCGCTACTGTGTCA 540
Qy       555 CCTAAAATATACCCCTTATCAATCGCTTCTTTAAACTCATCTATATATAACATATTTCAT 614
Db       541 CCTAAAATATACCCCTTATCAATCGCTTCTTTAAACTCATCTATATATAACATATTTCAT 600
Qy       615 CCTCCTACCTATCTATTCGTAAAAAGATAAAAATAACTATTGTTTTTTTTGTTATTTTAT 674
Db       601 CCTCCTACCTATCTATTCGTAAAAAGATAAAAATAACTATTGTTTTTTTTGTTATTTTAT 660
Qy       675 AATAAAATTATTAATATAAGTTAATGTTTTTTAAAAATATACAATTTTATTCTATTTATA 734
Db       661 AATAAAATTATTAATATAAGTTAATGTTTTTTAAAAATATACAATTTTATTCTATTTATA 720
Qy       735 GTTAGCTATTTTTTTCATTGTTAGTAATATTGGTGAATTGTAATAACCTTTTTAAATCTAG 794
Db       721 GTTAGCTATTTTTTTCATTGTTAGTAATATTGGTGAATTGTAATAACCTTTTTAAATCTAG 780
Qy       795 AGGAGAACCCAGATATAAAATGGAGGAATATTAATGGAAAACAATAAAAAAGTATTGAAG 854
Db       781 AGGAGAACCCAGATATAAAATGGAGGAATATTAATGGAAAACAATAAAAAAGTATTGAAG 840
Qy       855 AAAATGGTATTTTTTGTTTTAGTGACATTTCTTGGACTAACAATCTCGCAAGAGGTATTT 914
Db       841 AAAATGGTATTTTTTGTTTTAGTGACATTTCTTGGACTAACAATCTCGCAAGAGGTATTT 900
Qy       915 GCTCAACAAGACCCCGATCCAAGCCAACTTCACAGATCTAGTTTAGTTAAAAACCTTCAA 974
Db       901 GCTCAACAAGACCCCGATCCAAGCCAACTTCACAGATCTAGTTTAGTTAAAAACCTTCAA 960
Qy       975 AATATATATTTTCTTTATGAGGGTGACCCTGTTACTCACGAGAATGTGAAATCTGTTGAT 1034
Db       961 AATATATATTTTCTTTATGAGGGTGACCCTGTTACTCACGAGAATGTGAAATCTGTTGAT 1020
Qy      1035 CAACTTTTATCTCACCATTTAATATATAATGTTTCAGGGCCAAATTATGATAAATTAAAA 1094
Db      1021 CAACTTAGATCTCACGATTTAATATATAATGTTTCAGGGCCAAATTATGATAAATTAAAA 1080
Qy      1095 ACTGAACTTAAGAACCAAGAGATGGCAACTTTATTTAAGGATAAAAACGTTGATATTTAT 1154
Db      1081 ACTGAACTTAAGAACCAAGAGATGGCAACTTTATTTAAGGATAAAAACGTTGATATTTAT 1140
Qy      1155 GGTGTAGAATATTACCATCTCTGTTATTTATGTGAAAATGCAGAAAGGAGTGCATGTATC 1214
Db      1141 GGTGTAGAATATTACCATCTCTGTTATTTATGTGAAAATGCAGAAAGGAGTGCATGTATC 1200
Qy      1215 TACGGAGGGGTAACAAATCATGAAGGGAATCATTTAGAAATTCCTAAAAAGATAGTCGTT 1274
Db      1201 TACGGAGGGGTAACAAATCATGAAGGGAATCATTTAGAAATTCCTAAAAAGATAGTCGTT 1260
Qy      1275 AAAGTATCAATCGATGGTATCCAAAGCCTATCATTTGATATTGAAACAAATAAAAAAATG 1334
Db      1261 AAAGTATCAATCGATGGTATCCAAAGCCTATCATTTGATATTGAAACAAATAAAAAAATG 1320
Qy      1335 GTAACTGCTCAAGAATTAGACTATAAAGTTAGAAAATATCTTACAGATAATAAGCAACTA 1394
Db      1321 GTAACTGCTCAAGAATTAGACTATAAAGTTAGAAAATATCTTACAGATAATAAGCAACTA 1380
Qy      1395 TATACTAATGGACCTTCTAAATATGAAACTGGATATATAAAGTTCATACCTAAGAATAAA 1454
Db      1381 TATACTAATGGACCTTCTAAATATGAAACTGGATATATAAAGTTCATACCTAAGAATAAA 1440
Qy      1455 GAAAGTTTTTGGTTTGATTTTTTCCCTGAACCAGAATTTACTCAATCTAAATATCTTATG 1514
Db      1441 GAAAGTTTTTGGTTTGATTTTTTCCCTGAACCAGAATTTACTCAATCTAAATATCTTATG 1500
Qy      1515 ATATATAAAGATAATGAAACGCTTGACTCAAACACAAGCCAAATTGAAGTCTACCTAACA 1574
Db      1501 ATATATAAAGATAATGAAACGCTTGACTCAAACACAAGCCAAATTGAAGTCTACCTAACA 1560
Qy      1575 ACCAAGTAACTTTTTGCTTTTGGCAACCTTACCTACTGCTGGATTTAGAAATTTTATTGC 1634
Db      1561 ACCAAGTAACTTTTTGCTTTTGGCAACCTTACCTACTGCTGGATTTAGAAATTTTATTGC 1620
Qy      1635 AATTCTTTTATTAATGTAAAAACCGCTCATTTGATGAGCGGTTTTGTCTTATCTAAAGGA 1694
Db      1621 AATTCTTTTATTAATGTAAAAACCGCTCATTTGATGAGCGGTTTTGTCTTATCTAAAGGA 1680
Qy      1695 GCTTTACCTCCTAATGCTGCAAAATTTTAAATGTTGGATTTTTGTATTTGTCTATTGTAT 1754
Db      1681 GCTTTACCTCCTAATGCTGCAAAATTTTAAATGTTGGATTTTTGTATTTGTCTATTGTAT 1740
Qy      1755 TTGATGGGTAATCCCATTTTTCGACAGACATCGTCGTGCCACCTCTAACACCAAAATCAT 1814
Db      1741 TTGATGGGTAATCCCATTTTTCGACAGACATCGTCGTGCCACCTCTAACACCAAAATCAT 1800
Qy      1815 AGACAGGAGCTTGTAGCTTAGCAACTATTTTATCGTC 1851
Db      1801 AGACAGGAGCTTGTAGCTTAGCAACTATTTTATCGTC 1837
```

RESULT 8
AAD56771
ID AAD56771 standard; DNA; 1837 BP.
XX
AC AAD56771;
XX
DT 06-OCT-2003 (first entry)
XX
DE Streptococcus pyogenes pyrogenic exotoxin A (SpeA) DNA.
XX
KW Superantigen toxin; vaccine; infection; pyrogenic exotoxin A; SpeA;
KW gene therapy; gene; ds.
XX
OS Streptococcus pyogenes.
XX
FH Key   Location/Qualifiers
FT CDS   814. .1569
FT   /*tag= a
FT   /product= "Streptococcus pyogenes SpeA protein"
FT   /transl_except= (pos:901. .903, aa:Met)
FT   /note= "This translational exception occurs only when
FT   decoding SEQ ID NO: 26 (AAE37687)"
FT sig_peptide   814. .900
FT   /*tag= b
FT mat_peptide   901. .1566
FT   /*tag= c
FT   /product= "Streptococcus pyogenes mature SpeAprotein"
XX
PN WO200305601S-A1.
XX
PD 10-JUL-2003.
XX
PF 26-NOV-2001; 2001WO-US046540.
XX

PR 26-NOV-2001; 2001US-00002784.
XX
PA (USME-) US MEDICAL RES INST INFECTIOUS DISEASES.
XX
PI Ulrich RG;
XX
DR WPI; 2003-492125/46.
DR P-PSDB; AAE37683, AAE37687.
XX
PT New superantigen toxin DNA fragment, useful for preparing a composition
PT for treating or preventing bacterial infection.
XX
PS Claim 4; Page 128-129; 141pp; English.
XX
CC The invention provides a superantigen toxin DNA fragment which has been
CC genetically altered such that binding of the encoded altered toxin to
CC either the MHC class II or T cell antigen receptor is disrupted or
CC altered. DNA fragments of the invention are useful in the production of
CC vaccines against bacterial superantigen toxin infections. They are also
CC useful in gene therapy. The present sequence is Streptococcus pyogenes
CC pyrogenic exotoxin A (SpeA) DNA. This sequence is used in the invention
XX
SQ Sequence 1837 BP; 632 A; 318 C; 249 G; 638 T; 0 U; 0 Other;

Query Match 99.0%; Score 1832.2; DB 8; Length 1837;
Best Local Similarity 99.8%; Pred. No. 4.7e-286;
Matches 1834; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

```
Qy      15 TCATGTTTGACAGCTTATCATCGATAAGCTTACTTTTCGAATCAGGTCTATCCTTGAAAC 74
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1 TCATGTTTGACAGCTTATCATCGATAAGCTTACTTTTCGAATCAGGTCTATCCTTGAAAC 60

Qy      75 AGGTGCAACATAGATTAGGGCATGGAGATTTACCAGACAACTATGAACGTATATACTCAC 134
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      61 AGGTGCAACATAGATTAGGGCATGGAGATTTACCAGACAACTATGAACGTATATACTCAC 120

Qy     135 ATCACGCAATCGGCAATTGATGACATTGGAACTAAATTCAATCAATTTGTTACTAACAAG 194
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     121 ATCACGCAATCGGCAATTGATGACATTGGAACTAAATTCAATCAATTTGTTACTAACAAG 180

Qy     195 CAACTAGATTGACAACTAATTCTCAACAAACGTTAATTTAACAACATTCAAGTAACTCCC 254
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     181 CAACTAGATTGACAACTAATTCTCAACAAACGTTAATTTAACAACATTCAAGTAACTCCC 240

Qy     255 ACCAGCTCCATCAATGCTTACCGTAAGTAATCATAACTTACTAAAACCTTGTTACATCAA 314
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     241 ACCAGCTCCATCAATGCTTACCGTAAGTAATCATAACTTACTAAAACCTTGTTACATCAA 300

Qy     315 GGTTTTTTCTTTTTGTCTTGTTCATGAGTTACCATAACTTTCTATATTATTGACAACTAA 374
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     301 GGTTTTTTCTTTTTGTCTTGTTCATGAGTTACCATAACTTTCTATATTATTGACAACTAA 360

Qy     375 ATTGACAACTCTTCAATTATTTTTCTGTCTACTCAAAGTTTTCTTCATTTGATATAGTCT 434
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     361 ATTGACAACTCTTCAATTATTTTTCTGTCTACTCAAAGTTTTCTTCATTTGATATAGTCT 420

Qy     435 AATTCCACCATCACTTCTTCCACTCTCTCTACCGTCACAACTTCATCATCTCTCACTTTT 494
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     421 AATTCCACCATCACTTCTTCCACTCTCTCTACCGTCACAACTTCATCATCTCTCACTTTT 480

Qy     495 TCGTGTGGTAACACATAATCAAATATCTTTCCGTTTTTACGCACTATCGCTACTGTGTCA 554
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     481 TCGTGTGGTAACACATAATCAAATATCTTTCCGTTTTTACGCACTATCGCTACTGTGTCA 540

Qy     555 CCTAAAATATACCCCTTATCAATCGCTTCTTTAAACTCATCTATATATAACATATTTCAT 614
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     541 CCTAAAATATACCCCTTATCAATCGCTTCTTTAAACTCATCTATATATAACATATTTCAT 600

Qy     615 CCTCCTACCTATCTATTCGTAAAAAGATAAAAATAACTATTGTTTTTTTTGTTATTTTAT 674
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     601 CCTCCTACCTATCTATTCGTAAAAAGATAAAAATAACTATTGTTTTTTTTGTTATTTTAT 660

Qy     675 AATAAAATTATTAATATAAGTTAATGTTTTTTAAAAATATACAATTTTATTCTATTTATA 734
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     661 AATAAAATTATTAATATAAGTTAATGTTTTTTAAAAATATACAATTTTATTCTATTTATA 720

Qy     735 GTTAGCTATTTTTTCATTGTTAGTAATATTGGTGAATTGTAATAACCTTTTTAAATCTAG 794
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     721 GTTAGCTATTTTTTCATTGTTAGTAATATTGGTGAATTGTAATAACCTTTTTAAATCTAG 780

Qy     795 AGGAGAACCCAGATATAAAATGGAGGAATATTAATGGAAAACAATAAAAAAGTATTGAAG 854
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     781 AGGAGAACCCAGATATAAAATGGAGGAATATTAATGGAAAACAATAAAAAAGTATTGAAG 840

Qy     855 AAAATGGTATTTTTTGTTTTAGTGACATTTCTTGGACTAACAATCTCGCAAGAGGTATTT 914
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     841 AAAATGGTATTTTTTGTTTTAGTGACATTTCTTGGACTAACAATCTCGCAAGAGGTATTT 900

Qy     915 GCTCAACAAGACCCCGATCCAAGCCAACTTCACAGATCTAGTTTAGTTAAAAACCTTCAA 974
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     901 GCTCAACAAGACCCCGATCCAAGCCAACTTCACAGATCTAGTTTAGTTAAAAACCTTCAA 960

Qy     975 AATATATATTTTCTTTATGAGGGTGACCCTGTTACTCACGAGAATGTGAAATCTGTTGAT 1034
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     961 AATATATATTTTCTTTATGAGGGTGACCCTGTTACTCACGAGAATGTGAAATCTGTTGAT 1020

Qy    1035 CAACTTTTATCTCACCATTTAATATATAATGTTTCAGGGCCAAATTATGATAAATTAAAA 1094
           |||||| ||||||| ||||||||||||||||||||||||||||||||||||||||||||
Db    1021 CAACTTAGATCTCACGATTTAATATATAATGTTTCAGGGCCAAATTATGATAAATTAAAA 1080

Qy    1095 ACTGAACTTAAGAACCAAGAGATGGCAACTTTATTTAAGGATAAAAACGTTGATATTTAT 1154
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1081 ACTGAACTTAAGAACCAAGAGATGGCAACTTTATTTAAGGATAAAAACGTTGATATTTAT 1140

Qy    1155 GGTGTAGAATATTACCATCTCTGTTATTTATGTGAAAATGCAGAAAGGAGTGCATGTATC 1214
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1141 GGTGTAGAATATTACCATCTCTGTTATTTATGTGAAAATGCAGAAAGGAGTGCATGTATC 1200

Qy    1215 TACGGAGGGGTAACAAATCATGAAGGGAATCATTTAGAAATTCCTAAAAAGATAGTCGTT 1274
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1201 TACGGAGGGGTAACAAATCATGAAGGGAATCATTTAGAAATTCCTAAAAAGATAGTCGTT 1260

Qy    1275 AAAGTATCAATCGATGGTATCCAAAGCCTATCATTTGATATTGAAACAAATAAAAAAATG 1334
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1261 AAAGTATCAATCGATGGTATCCAAAGCCTATCATTTGATATTGAAACAAATAAAAAAATG 1320

Qy    1335 GTAACTGCTCAAGAATTAGACTATAAAGTTAGAAAATATCTTACAGATAATAAGCAACTA 1394
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1321 GTAACTGCTCAAGAATTAGACTATAAAGTTAGAAAATATCTTACAGATAATAAGCAACTA 1380

Qy    1395 TATACTAATGGACCTTCTAAATATGAAACTGGATATATATAAAGTTCATACCTAAGAATAAA 1454
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1381 TATACTAATGGACCTTCTAAATATGAAACTGGATATATATAAAGTTCATACCTAAGAATAAA 1440

Qy    1455 GAAAGTTTTTGGTTTGATTTTTTCCCTGAACCAGAATTTACTCAATCTAAATATCTTATG 1514
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1441 GAAAGTTTTTGGTTTGATTTTTTCCCTGAACCAGAATTTACTCAATCTAAATATCTTATG 1500

Qy    1515 ATATATAAAGATAATGAAACGCTTGACTCAAACACAAGCCAAATTGAAGTCTACCTAACA 1574
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1501 ATATATAAAGATAATGAAACGCTTGACTCAAACACAAGCCAAATTGAAGTCTACCTAACA 1560

Qy    1575 ACCAAGTAACTTTTTGCTTTTGGCAACCTTACCTACTGCTGGATTTAGAAATTTTATTGC 1634
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1561 ACCAAGTAACTTTTTGCTTTTGGCAACCTTACCTACTGCTGGATTTAGAAATTTTATTGC 1620

Qy    1635 AATTCTTTTATTAATGTAAAAACCGCTCATTTGATGAGCGGTTTTGTCTTATCTAAAGGA 1694
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1621 AATTCTTTTATTAATGTAAAAACCGCTCATTTGATGAGCGGTTTTGTCTTATCTAAAGGA 1680

Qy    1695 GCTTTACCTCCTAATGCTGCAAAATTTTAAATGTTGGATTTTTGTATTTGTCTATTGTAT 1754
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1681 GCTTTACCTCCTAATGCTGCAAAATTTTAAATGTTGGATTTTTGTATTTGTCTATTGTAT 1740

Qy    1755 TTGATGGGTAATCCCATTTTTCGACAGACATCGTCGTGCCACCTCTAACACCAAAATCAT 1814
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
```

```
Db       1741 TTGATGGGTAATCCCATTTTTCGACAGACATCGTCGTGCCACCTCTAACACCAAAATCAT 1800

Qy       1815 AGACAGGAGCTTGTAGCTTAGCAACTATTTTATCGTC 1851
              |||||||||||||||||||||||||||||||||||||
Db       1801 AGACAGGAGCTTGTAGCTTAGCAACTATTTTATCGTC 1837


RESULT 9
ACD28901
ID   ACD28901 standard; DNA; 1837 BP.
XX
AC   ACD28901;
XX
DT   27-AUG-2003  (first entry)
XX
DE   Streptococcus pyrogenic toxin a L42R mutant, DNA.
XX
KW   SPEa; streptococcus pyrogenic enterotoxin a; gene; mutant; vaccine; ds;
KW   superantigen toxin; MHC; superantigen-associated bacterial infection;
KW   bacterial infection; antibacterial.
XX
OS   Streptococcus sp.
OS   Synthetic.
XX
FH   Key             Location/Qualifiers
FT   CDS             814. .1569
FT                   /*tag=  a
FT                   /product= "SPEa L42R"
FT   sig_peptide     814. .903
FT                   /*tag=  b
FT   mat_peptide     904. .1566
FT                   /*tag=  c
FT                   /label= Mature_SPEa_L42R
XX
PN   US2003036644-A1.
XX
PD   20-FEB-2003.
XX
PF   26-NOV-2001; 2001US-00002784.
XX
PR   25-JUN-1997;   97US-00882431.
PR   01-SEP-1998;   98US-00144776.
XX
PA   (ULRI/) ULRICH R G.
XX
PI   Ulrich RG;
XX
DR   WPI; 2003-492125/46.
DR   P-PSDB; ABU62331.
XX
PT   New superantigen toxin DNA fragment, useful for preparing a composition
PT   for treating or preventing bacterial infection.
XX
PS   Claim 4; Page 36-37; 68pp; English.
XX
CC   The invention relates to an isolated and purified superantigen toxin DNA
CC   fragment is altered so that binding of the encoded altered toxin to
CC   either the MHC class II or T cell antigen receptor is altered. Also
CC   included are a recombinant DNA construct (comprising a vector and an
CC   isolated and purified altered superantigen toxin DNA fragment), a host
CC   cell transformed with the recombinant DNA construct, producing altered
CC   superantigen toxin, an altered TSST-1 (toxic shock syndrome toxin)
CC   superantigen toxin peptide, diagnosing superantigen-associated bacterial
CC   infection, a vaccine (comprising an altered superantigen toxin for
CC   producing antigenic and immunogenic response resulting in the protection
CC   of a mammal against superantigen-associated bacterial infection),
CC   treating/ameliorating a superantigen-associated bacterial infection, an
CC   antisera isolated from individuals immunised with one or more altered
CC   TSST-1 superantigen toxin and an antibody which recognises altered TSST-
CC   1. Also disclosed are mutated staphylococcal enterotoxin A, B, C1 (SEA,
CC   SEB, SEC1) and streptococcal pyrogenic enterotoxin a and b (SPEa and
CC   SPEb). The superantigen toxin DNA fragment is useful for preparing a
CC   composition for treating or preventing bacterial infection. The present
CC   sequence encodes the L42R (with reference to the mature protein) mutant
CC   of SPEa
XX
SQ   Sequence 1837 BP; 633 A; 318 C; 248 G; 638 T; 0 U; 0 Other;

  Query Match            98.9%;  Score 1830.6;  DB 8;  Length 1837;
  Best Local Similarity  99.8%;  Pred. No. 8.6e-286;
  Matches 1833;  Conservative    0;  Mismatches    4;  Indels    0;  Gaps    0;

Qy         15 TCATGTTTGACAGCTTATCATCGATAAGCTTACTTTTCGAATCAGGTCTATCCTTGAAAC 74
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 TCATGTTTGACAGCTTATCATCGATAAGCTTACTTTTCGAATCAGGTCTATCCTTGAAAC 60

Qy         75 AGGTGCAACATAGATTAGGGCATGGAGATTTACCAGACAACTATGAACGTATATACTCAC 134
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 AGGTGCAACATAGATTAGGGCATGGAGATTTACCAGACAACTATGAACGTATATACTCAC 120

Qy        135 ATCACGCAATCGGCAATTGATGACATTGGAACTAAATTCAATCAATTTGTTACTAACAAG 194
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 ATCACGCAATCGGCAATTGATGACATTGGAACTAAATTCAATCAATTTGTTACTAACAAG 180

Qy        195 CAACTAGATTGACAACTAATTCTCAACAAACGTTAATTTAACAACATTCAAGTAACTCCC 254
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 CAACTAGATTGACAACTAATTCTCAACAAACGTTAATTTAACAACATTCAAGTAACTCCC 240

Qy        255 ACCAGCTCCATCAATGCTTACCGTAAGTAATCATAACTTACTAAAACCTTGTTACATCAA 314
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 ACCAGCTCCATCAATGCTTACCGTAAGTAATCATAACTTACTAAAACCTTGTTACATCAA 300

Qy        315 GGTTTTTTCTTTTTGTCTTGTTCATGAGTTACCATAACTTTCTATATTATTGACAACTAA 374
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        301 GGTTTTTTCTTTTTGTCTTGTTCATGAGTTACCATAACTTTCTATATTATTGACAACTAA 360

Qy        375 ATTGACAACTCTTCAATTATTTTTCTGTCTACTCAAAGTTTTCTTCATTTGATATAGTCT 434
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        361 ATTGACAACTCTTCAATTATTTTTCTGTCTACTCAAAGTTTTCTTCATTTGATATAGTCT 420

Qy        435 AATTCCACCATCACTTCTTCCACTCTCTCTACCGTCACAACTTCATCATCTCTCACTTTT 494
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        421 AATTCCACCATCACTTCTTCCACTCTCTCTACCGTCACAACTTCATCATCTCTCACTTTT 480

Qy        495 TCGTGTGGTAACACATAATCAAATATCTTTCCGTTTTTACGCACTATCGCTACTGTGTCA 554
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        481 TCGTGTGGTAACACATAATCAAATATCTTTCCGTTTTTACGCACTATCGCTACTGTGTCA 540

Qy        555 CCTAAAATATACCCCTTATCAATCGCTTCTTTAAACTCATCTATATATAACATATTTCAT 614
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        541 CCTAAAATATACCCCTTATCAATCGCTTCTTTAAACTCATCTATATATAACATATTTCAT 600

Qy        615 CCTCCTACCTATCTATTCGTAAAAAGATAAAAATAACTATTGTTTTTTTTGTTATTTTAT 674
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        601 CCTCCTACCTATCTATTCGTAAAAAGATAAAAATAACTATTGTTTTTTTTGTTATTTTAT 660

Qy        675 AATAAAATTATTAATATAAGTTAATGTTTTTTAAAAATATACAATTTTATTCTATTTATA 734
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        661 AATAAAATTATTAATATAAGTTAATGTTTTTTAAAAATATACAATTTTATTCTATTTATA 720

Qy        735 GTTAGCTATTTTTTCATTGTTAGTAATATTGGTGAATTGTAATAACCTTTTTAAATCTAG 794
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        721 GTTAGCTATTTTTTCATTGTTAGTAATATTGGTGAATTGTAATAACCTTTTTAAATCTAG 780

Qy        795 AGGAGAACCCAGATATAAAATGGAGGAATATTAATGGAAAACAATAAAAAAGTATTGAAG 854
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        781 AGGAGAACCCAGATATAAAATGGAGGAATATTAATGGAAAACAATAAAAAAGTATTGAAG 840

Qy        855 AAAATGGTATTTTTTGTTTTAGTGACATTTCTTGGACTAACAATCTCGCAAGAGGTATTT 914
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        841 AAAATGGTATTTTTTGTTTTAGTGACATTTCTTGGACTAACAATCTCGCAAGAGGTATTT 900

Qy        915 GCTCAACAAGACCCCGATCCAAGCCAACTTCACAGATCTAGTTTAGTTAAAAACCTTCAA 974
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        901 GCTCAACAAGACCCCGATCCAAGCCAACTTCACAGATCTAGTTTAGTTAAAAACCTTCAA 960
```

```
Qy        975 AATATATATTTTCTTTATGAGGGTGACCCTGTTACTCACGAGAATGTGAAATCTGTTGAT 1034
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        961 AATATATATTTTCTTTATGAGGGTGACCCTGTTACTCACGAGAATGTGAAATCTGTTGAT 1020

Qy       1035 CAACTTTTATCTCACCATTTAATATATAATGTTTCAGGGCCAAATTATGATAAATTAAAA 1094
              ||||||  |||||||  ||||||||||||||||||||||||||||||||||||||||||||
Db       1021 CAACTTAGATCTCACGATTTAATATATAATGTTTCAGGGCCAAATTATGATAAATTAAAA 1080

Qy       1095 ACTGAACTTAAGAACCAAGAGATGGCAACTTTATTTAAGGATAAAAACGTTGATATTTAT 1154
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1081 ACTGAACTTAAGAACCAAGAGATGGCAACTTTATTTAAGGATAAAAACGTTGATATTTAT 1140

Qy       1155 GGTGTAGAATATTACCATCTCTGTTATTTATGTGAAAATGCAGAAAGGAGTGCATGTATC 1214
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1141 GGTGTAGAATATTACCATCTCTGTTATTTATGTGAAAATGCAGAAAGGAGTGCATGTATC 1200

Qy       1215 TACGGAGGGGTAACAAATCATGAAGGGAATCATTTAGAAATTCCTAAAAAGATAGTCGTT 1274
              ||||||||||||||||||||||||  ||||||||||||||||||||||||||||||||||
Db       1201 TACGGAGGGGTAACAAATCATGAAGGAAATCATTTAGAAATTCCTAAAAAGATAGTCGTT 1260

Qy       1275 AAAGTATCAATCGATGGTATCCAAAGCCTATCATTTGATATTGAAACAAATAAAAAAATG 1334
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1261 AAAGTATCAATCGATGGTATCCAAAGCCTATCATTTGATATTGAAACAAATAAAAAAATG 1320

Qy       1335 GTAACTGCTCAAGAATTAGACTATAAAGTTAGAAAATATCTTACAGATAATAAGCAACTA 1394
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1321 GTAACTGCTCAAGAATTAGACTATAAAGTTAGAAAATATCTTACAGATAATAAGCAACTA 1380

Qy       1395 TATACTAATGGACCTTCTAAATATGAAACTGGATATATAAAGTTCATACCTAAGAATAAA 1454
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1381 TATACTAATGGACCTTCTAAATATGAAACTGGATATATAAAGTTCATACCTAAGAATAAA 1440

Qy       1455 GAAAGTTTTTGGTTTGATTTTTTCCCTGAACCAGAATTTACTCAATCTAAATATCTTATG 1514
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1441 GAAAGTTTTTGGTTTGATTTTTTCCCTGAACCAGAATTTACTCAATCTAAATATCTTATG 1500

Qy       1515 ATATATAAAGATAATGAAACGCTTGACTCAAACACAAGCCAAATTGAAGTCTACCTAACA 1574
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1501 ATATATAAAGATAATGAAACGCTTGACTCAAACACAAGCCAAATTGAAGTCTACCTAACA 1560

Qy       1575 ACCAAGTAACTTTTTGCTTTTGGCAACCTTACCTACTGCTGGATTTAGAAATTTTATTGC 1634
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1561 ACCAAGTAACTTTTTGCTTTTGGCAACCTTACCTACTGCTGGATTTAGAAATTTTATTGC 1620

Qy       1635 AATTCTTTTATTAATGTAAAAACCGCTCATTTGATGAGCGGTTTTGTCTTATCTAAAGGA 1694
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1621 AATTCTTTTATTAATGTAAAAACCGCTCATTTGATGAGCGGTTTTGTCTTATCTAAAGGA 1680

Qy       1695 GCTTTACCTCCTAATGCTGCAAAATTTTAAATGTTGGATTTTTGTATTTGTCTATTGTAT 1754
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1681 GCTTTACCTCCTAATGCTGCAAAATTTTAAATGTTGGATTTTTGTATTTGTCTATTGTAT 1740

Qy       1755 TTGATGGGTAATCCCATTTTTCGACAGACATCGTCGTGCCACCTCTAACACCAAAATCAT 1814
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1741 TTGATGGGTAATCCCATTTTTCGACAGACATCGTCGTGCCACCTCTAACACCAAAATCAT 1800

Qy       1815 AGACAGGAGCTTGTAGCTTAGCAACTATTTTATCGTC 1851
              |||||||||||||||||||||||||||||||||||||
Db       1801 AGACAGGAGCTTGTAGCTTAGCAACTATTTTATCGTC 1837
```

RESULT 10
ACA64700
ID ACA64700 standard; DNA; 1031 BP.
XX
AC ACA64700;
XX
DT 18-JUN-2003 (first entry)
XX
DE S. pyogenes exotoxin (SPEA) gene.
XX
KW Superantigen; ds; gene; SAg; staphylococcal enterotoxin; tumour; cancer;
KW apoptosis; gene therapy; mammalian cell receptor; cytostatic;
KW tumour associated lipid; anergy; T cell; antigen presenting cell; APC;
KW tumouricidal immunocyte; antitumour.
XX
OS Streptococcus pyogenes.
XX
PN US2002177551-A1.
XX
PD 28-NOV-2002.
XX
PF 30-MAY-2001; 2001US-00870759.
XX
PR 31-MAY-2000; 2000US-0208128P.
XX
PA (TERM/) TERMAN D S.
XX
PI Terman DS;
XX
DR WPI; 2003-361759/34.
DR P-PSDB; ABU79074.
XX
PT A mammalian cell receptor, useful in the treatment of cancer by binding
PT to tumor associated lipids where the binding induces anergy or apoptosis
PT in T cells and antigen presenting cells.
XX
PS Disclosure; Page; 167pp; English.
XX
CC The invention relates to a mammalian cell receptor, useful in the
CC treatment of cancer, which binds to tumour associated lipids and induces
CC anergy or apoptosis in the T cells and antigen presenting cells (APCs).
CC Also included are a mammalian cell useful in the treatment of cancer
CC where the receptor which binds tumour associated lipids and induces
CC cellular inactivation or death is deleted or functionally deactivated,
CC producing (M1) a tumouricidal immunocyte population in vivo in a mammal
CC (by allowing tumour associated lipids to contact immunocytes in which
CC receptors for immunosuppressive fatty acids, ceramides, glycolipids,
CC sphingolipids, glycosphingolipids, phosphosphingolipids, gangliosides,
CC sialylated glycans, lipopeptides and proteoglycolipids are inactivated or
CC deleted), a construct useful in the treatment of cancer comprising a
CC superantigen (SAg) nucleotide inserted into a virus, a mammalian T cell
CC useful in the treatment of cancer (where an adaptor protein which
CC inhibits T cell activation by tumour associated antigens is deleted or
CC functionally deactivated), a composition useful in the treatment of
CC cancer (comprising a lipid raft conjugated to a superantigen), producing
CC (M2) a tumouricidal immunocyte population ex vivo in a mammal (by
CC allowing tumour associated lipids to contact immunocytes, in which
CC receptors for the lipids are inactivated or deleted to produce a
CC tumouricidal immunocyte population, and administering the tumouricidally
CC activated immunocytes to the host), producing (M3) a tumouricidal APC
CC population ex vivo in a mammal (by allowing a tumour associated lipid to
CC contact APCs, in which receptors for the tumour associated lipids are
CC inactivated or deleted to produce a tumouricidally activated population,
CC and administering APCs to the host), producing a tumouricidal T cell
CC population ex vivo in a mammal) by allowing a tumour associated lipids to
CC contact T cells, in which adaptor proteins, which inhibit T cell
CC activation by tumour associated antigens, are deleted or functionally
CC deactivated to produce a tumouricidal population of T cells, and
CC administering the tumouricidally activated T cells to the host, or
CC allowing a superantigen-lipid raft to contact T cells ex vivo, and
CC administering the tumouricidally activated T cells to the host), treating
CC (M5) cancer in a mammal (by administering a lipid binding molecule which
CC binds immunosuppressive tumour associated lipids in vivo), producing (M6)
CC a tumouricidal T cell population in vivo in a mammal (by allowing a
CC tumour associated antigen to contact immunocytes in which adaptor
CC proteins which inhibit T cell activation by tumour associated antigens
CC are deleted or functionally deactivated) and producing (M7) a
CC tumouricidal T cell population ex vivo in a mammal comprising allowing a
CC superantigen-lipid raft conjugate to contact immunocytes in vivo. The
CC receptors, methods and compositions are useful for treating cancers and
CC tumours. Bacterial superantigens are co-administered or administered as
CC fusion constructs with anti-tumour proteins or motifs. The present
CC sequence encodes a bacterial superantigen protein (e.g. a staphylococcal
CC enterotoxin). Note: The sequence data for this patent did not form part
CC of the printed specification, but was obtained in electronic format from

```
CC   the US patnet office website at
CC   "seqdata.uspto.gov/sequence.html?DocID=20020177551"
XX
SQ   Sequence 1031 BP; 381 A; 146 C; 162 G; 342 T; 0 U; 0 Other;

  Query Match             48.2%;  Score 892.8;  DB 7;  Length 1031;
  Best Local Similarity   95.7%;  Pred. No. 7.2e-135;
  Matches  993;  Conservative    0;  Mismatches   37;  Indels    8;  Gaps    7;

Qy        750 ATTGTTAGTAATATTGGTGAATTGTAATAACCTTTTTAAATCTAGAGGAGAACCCAGATA 809
              ||||||| ||||| |||| |||||||||||||||||||||||| ||||||||||||||||
Db          1 ATTGTTAATAATAATGGTTAATTGTAATAACCTTTTTAAATCTAAAGGAGAACCCAGATA 60

Qy        810 TAAAATGGAGGAATATTAATGGAAAACAATAAAAAAGTATTGAAGAAAATGGTATTTTTT 869
              |||||||||||||||||||||||||||||||| |||||||||||||||||||||||||||
Db         61 TAAAATGGAGGAATATTAATGGAAAACAATAAAGAAGTATTGAAGAAAATGGTATTTTTT 120

Qy        870 GTTTTAGTGACATTTCTTGGACTAACAATCTCGCAAGAGGTATTTGCTCAACAAGACCCC 929
              |||||| ||| |||||||||||||||||| || | | |   ||  |  |   |||
Db        121 GTTTTAATGAAATTTCTTGGACTAACAATCTTGCCAAAAGGTATTTGTTCAACAAGACCC 180

Qy        930 GATCCAAGCCAACTTCACAGATCTAGTTTAGTTAAAAACCTTCAAAATATATATTTTCTT 989
              | ||.|||||| | || |||||||| ||||||| ||||||||||||||||||||||||||
Db        181 AAGCCCAGCCAATTACAAAGATCTAATTTAGTT-AAAACCTTCAAAATATATATTTTCTT 239

Qy        990 TATGAGGGTGACCCT-GTTACTCACGAGAATGTGAAATCTGTTGATCAACTTTTATCTCA 1048
              |||||||||||||| |||||||||||||||||||||||||||||||||||||||||| ||
Db        240 TATGAGGGTGACCCTGGTTACTCACGAGAATGTGAAATCTGTTGATCAACTTTTATCACA 299

Qy       1049 CCATTTAATATATAATGTTTCAGGGCCAAATTATGATAAATTAAAAACTGAACTTAAGAA 1108
              | |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        300 CGATTTAATATATAATGTTTCAGGGCCAAATTATGATAAATTAAAAACTGAACTTAAGAA 359

Qy       1109 CCAAGAGATGGCAACTTTATTTAAGGATAAAAACGTTGATATTTATGGTGTAGAATATTA 1168
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        360 CCAAGAGATGGCAACTTTATTTAAGGATAAAAACGTTGATATTTATGGTGTAGAATATTA 419

Qy       1169 CCATCTCTGTTATTTATGTGAAAATGCAGAAAGGAGTGCATGTATCTACGGAGGGGTAAC 1228
              |||||||||||||||||||||||||||||||||||||||| ||||||||||||||||||
Db        420 CCATCTCTGTTATTTATGTGAAAATGCAGAAAGGAGTGCATGTCTCTACGGAGGGGTAAC 479

Qy       1229 AAATCATGAAGGGAATCATTTAGAAATTCCTAAAAAGATAGTCGTTAAAGTATCAATCGA 1288
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        480 AAATCATGAAGGGAATCATTTAGAAATTCCTAAAAAGATAGTCGTTAAAGTATCAATCGA 539

Qy       1289 TGGTATCCAAAGCCTATCATTTGATATTGAAACAAATAAAAAAAATGGTAACTGCTCAAGA 1348
              |||||||||||||||||||||||||||||||    ||  ||||||||||||||||||||
Db        540 TGGTATCCAAAGCCTATCATTTGATATTGAA--CAAATAAAAAATGGTAACTGCTCAAGA 597

Qy       1349 ATTAGACTATAAAGTTAGAAAATATCTTACAGATAATAAGCAACTATATACTAATGGACC 1408
              |||||  |||||  |||||||||||||||||||||||||||||||||||||||||||||
Db        598 ATTAG-CTATACAGTTAGAAAATATCTTACAGATAATAAGCAACTATATACTAATGGACC 656

Qy       1409 TTCTAAATATGAAACTGGATATATAAAGTTCATACCTAAGAATAAAGAAAGTTTTTGGTT 1468
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        657 TTCTAAATATGAAACTGGATATATAAAGTTCATACCTAAGAATAAAGAAAGTTTTTGGTT 716

Qy       1469 TGATTTTTTCCCTGAACCAGAATTTACTCAATCTAAATATCTTATGATATATAAAGATAA 1528
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        717 TGATTTTTTCCCTGAACCAGAATTTACTCAATCTAAATATCTTATGATATATAAAGATAA 776

Qy       1529 TGAAACGCTTGACTCAAACACAAGCCAAATTGAAGTCTACCTAACAACCAAGTAACTTTT 1588
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||| |||
Db        777 TGAAACGCTTGACTCAAACACAAGCCAAATTGAAGTCTACCTAACAACCAAGTAAC-TTT 835

Qy       1589 TGCTTTTGGCAACCTTACCTACTGCTGGATTTAGAAATTTTATTGCAATTCTTTTATTAA 1648
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        836 TGCTTTTGGCAACCTTACCTACTGCTGGATTTAGAAATTTTATTGCAATTCTTTTATTAA 895

Qy       1649 TGTAAAAACCGCTCATTTGATGAGCGGTTTTGTCTTATCTAAAGGAGCTTTACCTCCTAA 1708
              |||||||| |||||||||||||||||||||||||||||||||||||||||||||||||||
Db        896 TGTAAAAA-CGCTCATTTGATGAGCGGTTTTGTCTTATCTAAAGGAGCTTTACCTCCTAA 954

Qy       1709 TGCTGCAAAATTTTAAATGTTGGATTTTTGTATTTGTCTATTGTATTTGATGGGTAATCC 1768
              |||||||||||||||||||||||||||||||||||||||| |||||||||||||||||||
Db         955 TGCTGCAAAATTTTAAATGTTGGATTTTTGTATTTGTC-ATTGTATTTGATGGGTAATCC 1013

Qy       1769 CATTTTTCGACAGACATC 1786
              ||||||||||||||||||
Db       1014 CATTTTTCGACAGACATC 1031


RESULT 11
AAH01002
ID   AAH01002 standard; DNA; 756 BP.
XX
AC   AAH01002;
XX
DT   24-JUL-2001  (first entry)
XX
DE   Unidentified nucleotide sequence SEQ ID NO:993.
XX
KW   Species specific; genus specific; family specific; probe; detection;
KW   identification; algal; archaeal; bacterial; fungal; parasitical;
KW   microorganism; diagnosis; translation elongation factor Tu; toxin;
KW   translation elongation factor G; RecA recombinase; resistance;
KW   catalytic subunit of proton-translocating ATPase; antimicrobial; vaccine;
KW   primer; ds.
XX
OS   Unidentified.
XX
PN   WO200123604-A2.
XX
PD   05-APR-2001.
XX
PF   28-SEP-2000; 2000WO-CA001150.
XX
PR   28-SEP-1999;   99CA-02283458.
PR   19-MAY-2000; 2000CA-02307010.
XX
PA   (INPE-) INFECTIO DIAGNOSTIC (IDI) INC.
XX
PI   Bergeron MG,  Boissinot M,  Huletsky A,  Menard C,  Ouellette M;
PI   Picard FJ,  Roy PH;
XX
DR   WPI; 2001-245006/25.
XX
PT   Nucleic acid sequences are used to generate universal probes and primers
PT   which can be used to identify and detect the presence of algal, archaeal,
PT   bacterial, fungal and parasitical species in a test sample.
XX
PS   Disclosure; Page 957; 1580pp; English.
XX
CC   The present invention describes a method for generating a repertory of
CC   nucleic acids of tuf, fus, atpD and/or recA genes from which probes
CC   and/or primers are derived. The method comprises amplifying the nucleic
CC   acids of determined algal, archaeal, bacterial, fungal and parasitical
CC   species with a combination of defined primer pairs. The method can be
CC   used for producing probes and/or primers for detecting one or more
CC   related microorganisms e.g. algae, archaea, bacteria, fungi and
CC   parasites, for universal detection and for specific and ubiquitous
CC   detection and identification of an algal, archaeal, bacterial, fungal and
CC   parasitical species, genus, family and group. A nucleic acid (I) obtained
CC   using the method of the invention can be used for the universal detection
CC   of any bacterium, fungus or parasite in a sample and for the detection of
CC   at least one antimicrobial agent resistance gene or at least one toxin
CC   gene. hexA nucleic acids are used for the specific and ubiquitous
CC   detection and for identification of Streptococcus pneumoniae. (I) can be
CC   used to design a therapeutic agent which is effective against
CC   microorganisms. Microbial species or genus or family or phylum or group
CC   which can be detected include Abiotrophia adiacens, Bordetella sp.,
CC   Corynebacterium sp., Enterobacteriaceae group, Escherichia coli,
CC   Mycobacteriaceae family, Pseudomonads group, Streptococcus sp., Neisseria
CC   gonorrhoeae and Staphylococcus sp. . Using DNA based tests provides faster
CC   results than substrate specificity tests as results can be determined in
CC   an hour and improved accuracy is also achieved. AAH00010 to AAH002304
```

```
CC   represent nucleotide sequences and primers/probes which are given in the
CC   exemplification of the present invention
XX
SQ   Sequence 756 BP; 297 A; 110 C; 118 G; 231 T; 0 U; 0 Other;

  Query Match             40.8%;  Score 754.4;  DB 4;  Length 756;
  Best Local Similarity   99.9%;  Pred. No. 1.4e-112;
  Matches  755;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy        828 ATGGAAAACAATAAAAAAGTATTGAAGAAAATGGTATTTTTTGTTTTAGTGACATTTCTT 887
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 ATGGAAAACAATAAAAAAGTATTGAAGAAAATGGTATTTTTTGTTTTAGTGACATTTCTT 60

Qy        888 GGACTAACAATCTCGCAAGAGGTATTTGCTCAACAAGACCCCGATCCAAGCCAACTTCAC 947
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 GGACTAACAATCTCGCAAGAGGTATTTGCTCAACAAGACCCCGATCCAAGCCAACTTCAC 120

Qy        948 AGATCTAGTTTAGTTAAAAACCTTCAAAATATATATTTTCTTTATGAGGGTGACCCTGTT 1007
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 AGATCTAGTTTAGTTAAAAACCTTCAAAATATATATTTTCTTTATGAGGGTGACCCTGTT 180

Qy       1008 ACTCACGAGAATGTGAAATCTGTTGATCAACTTTTATCTCACCATTTAATATATAATGTT 1067
              |||||||||||||||||||||||||||||||||||||||||| ||||||||||||||||
Db        181 ACTCACGAGAATGTGAAATCTGTTGATCAACTTTTATCTCACGATTTAATATATAATGTT 240

Qy       1068 TCAGGGCCAAATTATGATAAATTAAAAACTGAACTTAAGAACCAAGAGATGGCAACTTTA 1127
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 TCAGGGCCAAATTATGATAAATTAAAAACTGAACTTAAGAACCAAGAGATGGCAACTTTA 300

Qy       1128 TTTAAGGATAAAAACGTTGATATTTATGGTGTAGAATATTACCATCTCTGTTATTTATGT 1187
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        301 TTTAAGGATAAAAACGTTGATATTTATGGTGTAGAATATTACCATCTCTGTTATTTATGT 360

Qy       1188 GAAAATGCAGAAAGGAGTGCATGTATCTACGGAGGGGTAACAAATCATGAAGGGAATCAT 1247
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        361 GAAAATGCAGAAAGGAGTGCATGTATCTACGGAGGGGTAACAAATCATGAAGGGAATCAT 420

Qy       1248 TTAGAAATTCCTAAAAAGATAGTCGTTAAAGTATCAATCGATGGTATCCAAAGCCTATCA 1307
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        421 TTAGAAATTCCTAAAAAGATAGTCGTTAAAGTATCAATCGATGGTATCCAAAGCCTATCA 480

Qy       1308 TTTGATATTGAAACAAATAAAAAAATGGTAACTGCTCAAGAATTAGACTATAAAGTTAGA 1367
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        481 TTTGATATTGAAACAAATAAAAAAATGGTAACTGCTCAAGAATTAGACTATAAAGTTAGA 540

Qy       1368 AAATATCTTACAGATAATAAGCAACTATATACTAATGGACCTTCTAAATATGAAACTGGA 1427
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        541 AAATATCTTACAGATAATAAGCAACTATATACTAATGGACCTTCTAAATATGAAACTGGA 600

Qy       1428 TATATAAAGTTCATACCTAAGAATAAAGAAAGTTTTTGGTTTGATTTTTTCCCTGAACCA 1487
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        601 TATATAAAGTTCATACCTAAGAATAAAGAAAGTTTTTGGTTTGATTTTTTCCCTGAACCA 660

Qy       1488 GAATTTACTCAATCTAAATATCTTATGATATATAAAGATAATGAAACGCTTGACTCAAAC 1547
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        661 GAATTTACTCAATCTAAATATCTTATGATATATAAAGATAATGAAACGCTTGACTCAAAC 720

Qy       1548 ACAAGCCAAATTGAAGTCTACCTAACAACCAAGTAA 1583
              ||||||||||||||||||||||||||||||||||||
Db        721 ACAAGCCAAATTGAAGTCTACCTAACAACCAAGTAA 756


RESULT 12
ACD28908
ID   ACD28908 standard; DNA; 1419 BP.
XX
AC   ACD28908;
XX
DT   27-AUG-2003  (first entry)
XX
DE   SPEa L42R/SPEb C47S mutant fusion protein, DNA.
XX
KW   SPEa; streptococcus pyrogenic enterotoxin; gene; mutant; vaccine; ds;
KW   superantigen toxin; MHC; superantigen-associated bacterial infection;
KW   bacterial infection; antibacterial; SPEb.
XX
OS   Streptococcus sp.
OS   Synthetic.
XX
FH   Key             Location/Qualifiers
FT   CDS             1. .1419
FT                   /*tag=  a
FT                   /product= "SPEa L42R/SPEb C47S"
FT                   /transl_except= (pos:298. .306,aa:Ile-Gly)
FT                   /transl_except= (pos:634. .642,aa:Thr-Gln)
FT                   /transl_except= (pos:1162. .1170,aa:Ser-Gln)
FT                   /transl_except= (pos:1228. .1236,aa:Gly-Gly)
XX
PN   US2003036644-A1.
XX
PD   20-FEB-2003.
XX
PF   26-NOV-2001; 2001US-00002784.
XX
PR   25-JUN-1997;   97US-00882431.
PR   01-SEP-1998;   98US-00144776.
XX
PA   (ULRI/) ULRICH R G.
XX
PI   Ulrich RG;
XX
DR   WPI; 2003-492125/46.
DR   P-PSDB; ABU62335.
XX
PT   New superantigen toxin DNA fragment, useful for preparing a composition
PT   for treating or preventing bacterial infection.
XX
PS   Claim 6; Page 39-40; 68pp; English.
XX
CC   The invention relates to an isolated and purified superantigen toxin DNA
CC   fragment is altered so that binding of the encoded altered toxin to
CC   either the MHC class II or T cell antigen receptor is altered. Also
CC   included are a recombinant DNA construct (comprising a vector and an
CC   isolated and purified altered superantigen toxin DNA fragment), a host
CC   cell transformed with the recombinant DNA construct, producing altered
CC   superantigen toxin, an altered TSST-1 (toxic shock syndrome toxin)
CC   superantigen toxin peptide, diagnosing superantigen-associated bacterial
CC   infection, a vaccine (comprising an altered superantigen toxin for
CC   producing antigenic and immunogenic response resulting in the protection
CC   of a mammal against superantigen-associated bacterial infection),
CC   treating/ameliorating a superantigen-associated bacterial infection, an
CC   antisera isolated from individuals immunised with one or more altered
CC   TSST-1 superantigen toxin and an antibody which recognises altered TSST-
CC   1. Also disclosed are mutated staphylococcal enterotoxin A, B, C1 (SEA,
CC   SEB, SEC1) and streptococcal pyrogenic enterotoxin a and b (SPEa and
CC   SPEb). The superantigen toxin DNA fragment is useful for preparing a
CC   composition for treating or preventing bacterial infection. The present
CC   sequence encodes the SPEa L42R/SPEb C47S mutant fusion protein
XX
SQ   Sequence 1419 BP; 497 A; 266 C; 255 G; 401 T; 0 U; 0 Other;

  Query Match             35.5%;  Score 656.8;  DB 8;  Length 1419;
  Best Local Similarity   94.2%;  Pred. No. 7e-97;
  Matches  682;  Conservative    0;  Mismatches   42;  Indels    0;  Gaps    0;

Qy        918 CAACAAGACCCCGATCCAAGCCAACTTCACAGATCTAGTTTAGTTAAAAACCTTCAAAAT 977
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          4 CAACAAGACCCCGATCCAAGCCAACTTCACAGATCTAGTTTAGTTAAAAACCTTCAAAAT 63

Qy        978 ATATATTTTCTTTATGAGGGTGACCCTGTTACTCACGAGAATGTGAAATCTGTTGATCAA 1037
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         64 ATATATTTTCTTTATGAGGGTGACCCTGTTACTCACGAGAATGTGAAATCTGTTGATCAA 123

Qy       1038 CTTTTATCTCACCATTTAATATATAATGTTTCAGGGCCAAATTATGATAAATTAAAAACT 1097
              |||  |||||||  ||||||||||||||||||||||||||||||||||||||||||||||
Db        124 CTTCGATCTCACGATTTAATATATAATGTTTCAGGGCCAAATTATGATAAATTAAAAACT 183
```

```
Qy       1098 GAACTTAAGAACCAAGAGATGGCAACTTTATTTAAGGATAAAAACGTTGATATTTATGGT 1157
             ||||||||||||||||||||||||||||||||||||||||||||| ||||||||||||||
Db        184 GAACTTAAGAACCAAGAGATGGCAACTTTATTTAAGGATAAAAACATTGATATTTATGGT 243

Qy       1158 GTAGAATATTACCATCTCTGTTATTTATGTGAAAATGCAGAAAGGAGTGCATGTATCTAC 1217
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        244 GTAGAATATTACCATCTCTGTTATTTATGTGAAAATGCAGAAAGGAGTGCATGTATCTAC 303

Qy       1218 GGAGGGGTAACAAATCATGAAGGGAATCATTTAGAAATTCCTAAAAAGATAGTCGTTAAA 1277
             |||||||||||||||| |||||||||||||||||||||||||||||||||||||||| |||
Db        304 GGAGGGGTAACAAATCGTGAAGGGAATCATTTAGAAATTCCTAAAAAGATAGTCGTCAAA 363

Qy       1278 GTATCAATCGATGGTATCCAAAGCCTATCATTTGATATTGAAACAAATAAAAAAATGGTA 1337
             ||||||||||||||| |||||||||||||||||||||||||||||||||||||||||||||
Db        364 GTATCAATCGATGGTATACAAAGCCTATCATTTGATATTGAAACAAATAAAAAAATGGTA 423

Qy       1338 ACTGCTCAAGAATTAGACTATAAAGTTAGAAAATATCTTACAGATAATAAGCAACTATAT 1397
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        424 ACTGCTCAAGAATTAGACTATAAAGTTAGAAAATATCTTACAGATAATAAGCAACTATAT 483

Qy       1398 ACTAATGGACCTTCTAAATATGAAACTGGATATATAAAGTTCATACCTAAGAATAAAGAA 1457
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        484 ACTAATGGACCTTCTAAATATGAAACTGGATATATAAAGTTCATACCTAAGAATAAAGAA 543

Qy       1458 AGTTTTTGGTTTGATTTTTTCCCTGAACCAGAATTTACTCAATCTAAATATCTTATGATA 1517
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        544 AGTTTTTGGTTTGATTTTTTCCCTGAACCAGAATTTACTCAATCTAAATATCTTATGATA 603

Qy       1518 TATAAAGATAATGAAACGCTTGACTCAAACACAAGCCAAATTGAAGTCTACCTAACAACC 1577
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        604 TATAAAGATAATGAAACGCTTGACTCAAACACAAGCCAAATTGAAGTCTACCTAACAACC 663

Qy       1578 AAGTAACTTTTTGCTTTTGGCAACCTTACCTACTGCTGGATTTAGAAATTTTATTGCAAT 1637
             ||| ||| ||| |         |||| |       | ||| |  |  | | | | ||
Db        664 AAGCAACCAGTTGTTAAATCTCTCCTTGATTCAAAAGGCATTCATTACAATCAAGGTAAC 723

Qy       1638 TCTT 1641
                ||||
Db        724 CCTT 727
```

```
RESULT 13
AAD56778
ID   AAD56778 standard; DNA; 1419 BP.
XX
AC   AAD56778;
XX
DT   06-OCT-2003  (first entry)
XX
DE   Streptococcus pyogenes SpeA-SpeB fusion DNA.
XX
KW   Superantigen toxin; vaccine; infection; pyrogenic exotoxin A; SpeA;
KW   gene therapy; fusion protein; SpeB; gene; ds.
XX
OS   Streptococcus pyogenes.
XX
FH   Key             Location/Qualifiers
FT   CDS             1. .1419
FT                   /*tag= a
FT                   /product= "S. pyogenes SpeA-SpeB fusion protein"
FT                   /transl_except= (pos:298. .306, aa:Ile-Gly)
FT                   /transl_except= (pos:634. .642, aa:Thr-Gln)
FT                   /transl_except= (pos:1162. .1170, aa:Ser-Gln)
FT                   /transl_except= (pos:1228. .1236, aa:Gly-Gly)
XX
PN   WO2003056015-A1.
XX
PD   10-JUL-2003.
XX
PF   26-NOV-2001; 2001WO-US046540.
XX
PR   26-NOV-2001; 2001US-00002784.
XX
PA   (USME-) US MEDICAL RES INST INFECTIOUS DISEASES.
XX
PI   Ulrich RG;
XX
DR   WPI; 2003-492125/46.
DR   P-PSDB; AAE37684.
XX
PT   New superantigen toxin DNA fragment, useful for preparing a composition
PT   for treating or preventing bacterial infection.
XX
PS   Claim 6; Page 132-133; 141pp; English.
XX
CC   The invention provides a superantigen toxin DNA fragment which has been
CC   genetically altered such that binding of the encoded altered toxin to
CC   either the MHC class II or T cell antigen receptor is disrupted or
CC   altered. DNA fragments of the invention are useful in the production of
CC   vaccines against bacterial superantigen toxin infections. They are also
CC   useful in gene therapy. The present sequence is Streptococcus pyogenes
CC   pyrogenic exotoxin A (SpeA)-SpeB fusion DNA. This sequence is used in the
CC   invention
XX
SQ   Sequence 1419 BP; 497 A; 266 C; 255 G; 401 T; 0 U; 0 Other;
```

```
  Query Match           35.5%;  Score 656.8;  DB 8;  Length 1419;
  Best Local Similarity  94.2%;  Pred. No. 7e-97;
  Matches  682;  Conservative    0;  Mismatches   42;  Indels    0;  Gaps    0;

Qy        918 CAACAAGACCCCGATCCAAGCCAACTTCACAGATCTAGTTTAGTTAAAAACCTTCAAAAT 977
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          4 CAACAAGACCCCGATCCAAGCCAACTTCACAGATCTAGTTTAGTTAAAAACCTTCAAAAT 63

Qy        978 ATATATTTTCTTTATGAGGGTGACCCTGTTACTCACGAGAATGTGAAATCTGTTGATCAA 1037
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         64 ATATATTTTCTTTATGAGGGTGACCCTGTTACTCACGAGAATGTGAAATCTGTTGATCAA 123

Qy       1038 CTTTTATCTCACCATTTAATATATAATGTTTCAGGGCCAAATTATGATAAATTAAAAACT 1097
              ||| ||||||| ||||||||||||||||||||||||||||||||||||||||||||||||
Db        124 CTTCGATCTCACGATTTAATATATAATGTTTCAGGGCCAAATTATGATAAATTAAAAACT 183

Qy       1098 GAACTTAAGAACCAAGAGATGGCAACTTTATTTAAGGATAAAAACGTTGATATTTATGGT 1157
              ||||||||||||||||||||||||||||||||||||||||||||| ||||||||||||||
Db        184 GAACTTAAGAACCAAGAGATGGCAACTTTATTTAAGGATAAAAACATTGATATTTATGGT 243

Qy       1158 GTAGAATATTACCATCTCTGTTATTTATGTGAAAATGCAGAAAGGAGTGCATGTATCTAC 1217
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        244 GTAGAATATTACCATCTCTGTTATTTATGTGAAAATGCAGAAAGGAGTGCATGTATCTAC 303

Qy       1218 GGAGGGGTAACAAATCATGAAGGGAATCATTTAGAAATTCCTAAAAAGATAGTCGTTAAA 1277
              |||||||||||||||| |||||||||||||||||||||||||||||||||||||||| |||
Db        304 GGAGGGGTAACAAATCGTGAAGGGAATCATTTAGAAATTCCTAAAAAGATAGTCGTCAAA 363

Qy       1278 GTATCAATCGATGGTATCCAAAGCCTATCATTTGATATTGAAACAAATAAAAAAATGGTA 1337
              ||||||||||||||| |||||||||||||||||||||||||||||||||||||||||||||
Db        364 GTATCAATCGATGGTATACAAAGCCTATCATTTGATATTGAAACAAATAAAAAAATGGTA 423

Qy       1338 ACTGCTCAAGAATTAGACTATAAAGTTAGAAAATATCTTACAGATAATAAGCAACTATAT 1397
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        424 ACTGCTCAAGAATTAGACTATAAAGTTAGAAAATATCTTACAGATAATAAGCAACTATAT 483

Qy       1398 ACTAATGGACCTTCTAAATATGAAACTGGATATATAAAGTTCATACCTAAGAATAAAGAA 1457
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        484 ACTAATGGACCTTCTAAATATGAAACTGGATATATAAAGTTCATACCTAAGAATAAAGAA 543

Qy       1458 AGTTTTTGGTTTGATTTTTTCCCTGAACCAGAATTTACTCAATCTAAATATCTTATGATA 1517
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        544 AGTTTTTGGTTTGATTTTTTCCCTGAACCAGAATTTACTCAATCTAAATATCTTATGATA 603

Qy       1518 TATAAAGATAATGAAACGCTTGACTCAAACACAAGCCAAATTGAAGTCTACCTAACAACC 1577
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        604 TATAAAGATAATGAAACGCTTGACTCAAACACAAGCCAAATTGAAGTCTACCTAACAACC 663

Qy       1578 AAGTAACTTTTTGCTTTTGGCAACCTTACCTACTGCTGGATTTAGAAATTTTATTGCAAT 1637
```

```
                ||| |||   ||| |        ||||   |      | ||| |  |   | |  | ||
Db        664 AAGCAACCAGTTGTTAAATCTCTCCTTGATTCAAAAGGCATTCATTACAATCAAGGTAAC 723

Qy       1638 TCTT 1641
                |||
Db        724 CCTT 727
```

RESULT 14
ACA64721
ID ACA64721 standard; DNA; 1497 BP.
XX
AC ACA64721;
XX
DT 18-JUN-2003 (first entry)
XX
DE Staphylococcus/Streptococcus erythrogenic toxin gene.
XX
KW Superantigen; ds; gene; SAg; staphylococcal enterotoxin; tumour; cancer;
KW apoptosis; gene therapy; mammalian cell receptor; cytostatic;
KW tumour associated lipid; anergy; T cell; antigen presenting cell; APC;
KW tumouricidal immunocyte; antitumour.
XX
OS Staphylococcus sp.
OS Streptococcus sp.
XX
PN US20020177551-A1.
XX
PD 28-NOV-2002.
XX
PF 30-MAY-2001; 2001US-00870759.
XX
PR 31-MAY-2000; 2000US-0208128P.
XX
PA (TERM/) TERMAN D S.
XX
PI Terman DS;
XX
DR WPI; 2003-361759/34.
DR P-PSDB; ABU79113.
XX
PT A mammalian cell receptor, useful in the treatment of cancer by binding
PT to tumor associated lipids where the binding induces anergy or apoptosis
PT in T cells and antigen presenting cells.
XX
PS Example 2; Page; 167pp; English.
XX
CC The invention relates to a mammalian cell receptor, useful in the
CC treatment of cancer, which binds to tumour associated lipids and induces
CC anergy or apoptosis in the T cells and antigen presenting cells (APCs).
CC Also included are a mammalian cell useful in the treatment of cancer
CC where the receptor which binds tumour associated lipids and induces
CC cellular inactivation or death is deleted or functionally deactivated,
CC producing (M1) a tumouricidal immunocyte population in vivo in a mammal
CC (by allowing tumour associated lipids to contact immunocytes in which
CC receptors for immunosuppressive fatty acids, ceramides, glycolipids,
CC sphingolipids, glycosphingolipids, phosphosphingolipids, gangliosides,
CC sialylated glycans, lipopeptides and proteoglycolipids are inactivated or
CC deleted), a construct useful in the treatment of cancer comprising a
CC superantigen (SAg) nucleotide inserted into a virus, a mammalian T cell
CC useful in the treatment of cancer (where an adaptor protein which
CC inhibits T cell activation by tumour associated antigens is deleted or
CC functionally deactivated), a composition useful in the treatment of
CC cancer (comprising a lipid raft conjugated to a superantigen), producing
CC (M2) a tumouricidal immunocyte population ex vivo in a mammal (by
CC allowing tumour associated lipids to contact immunocytes, in which
CC receptors for the lipids are inactivated or deleted to produce a
CC tumouricidal immunocyte population, and administering the tumouricidally
CC activated immunocytes to the host), producing (M3) a tumouricidal APC
CC population ex vivo in a mammal (by allowing a tumour associated lipid to
CC contact APCs, in which receptors for the tumour associated lipids are
CC inactivated or deleted to produce a tumouricidally activated population,
CC and administering APCs to the host), producing a tumouricidal T cell
CC population ex vivo in a mammal) by allowing a tumour associated lipids to
CC contact T cells, in which adaptor proteins, which inhibit T cell
CC activation by tumour associated antigens, are deleted or functionally
CC deactivated to produce a tumouricidal population of T cells, and
CC administering the tumouricidally activated T cells to the host, or
CC allowing a superantigen-lipid raft to contact T cells ex vivo, and
CC administering the tumouricidally activated T cells to the host), treating
CC (M5) cancer in a mammal (by administering a lipid binding molecule which
CC binds immunosuppressive tumour associated lipids in vivo), producing (M6)
CC a tumouricidal T cell population in vivo in a mammal (by allowing a
CC tumour associated antigen to contact immunocytes in which adaptor
CC proteins which inhibit T cell activation by tumour associated antigens
CC are deleted or functionally deactivated) and producing (M7) a
CC tumouricidal T cell population ex vivo in a mammal comprising allowing a
CC superantigen-lipid raft conjugate to contact immunocytes in vivo. The
CC receptors, methods and compositions are useful for treating cancers and
CC tumours. Bacterial superantigens are co-administered or administered as
CC fusion constructs with anti-tumour proteins or motifs. The present
CC sequence encodes an anti-tumour protein which is co-administered with or
CC incorporated into a fusion construct with a superantigen. Note: The
CC sequence data for this patent did not form part of the printed
CC specification, but was obtained in electronic format from the US patnet
CC office website at "seqdata.uspto.gov/sequence.html?DocID=20020177551"
XX
SQ Sequence 1497 BP; 584 A; 247 C; 246 G; 420 T; 0 U; 0 Other;

  Query Match   15.0%; Score 278; DB 7; Length 1497;
  Best Local Similarity   99.7%; Pred. No. 6.5e-36;
  Matches 289; Conservative   0; Mismatches   0; Indels   1; Gaps   1;

```
Qy         40 AAGCTTACTTTTCGAATCAGGTCTATCCTTGAAACAGGTGCAACATAGATTAGGGCATGG 99
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1209 AAGCTTACTTTTCGAATCAGGTCTATCCTTGAAACAGGTGCAACATAGATTAGGGCATGG 1268

Qy        100 AGATTTACCAGACAACTATGAACGTATATACTCACATCACGCAATCGGCAATTGATGACA 159
              ||||||| ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1269 AGATTTA-CAGACAACTATGAACGTATATACTCACATCACGCAATCGGCAATTGATGACA 1327

Qy        160 TTGGAACTAAATTCAATCAATTTGTTACTAACAAGCAACTAGATTGACAACTAATTCTCA 219
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1328 TTGGAACTAAATTCAATCAATTTGTTACTAACAAGCAACTAGATTGACAACTAATTCTCA 1387

Qy        220 ACAAACGTTAATTTAACAACATTCAAGTAACTCCCACCAGCTCCATCAATGCTTACCGTA 279
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1388 ACAAACGTTAATTTAACAACATTCAAGTAACTCCCACCAGCTCCATCAATGCTTACCGTA 1447

Qy        280 AGTAATCATAACTTACTAAAACCTTGTTACATCAAGGTTTTTTCTTTTTG 329
              ||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1448 AGTAATCATAACTTACTAAAACCTTGTTACATCAAGGTTTTTTCTTTTTG 1497
```

RESULT 15
AAZ51107
ID AAZ51107 standard; DNA; 1712 BP.
XX
AC AAZ51107;
XX
DT 05-JUN-2000 (first entry)
XX
DE Staphylococcal enterotoxin B encoding DNA.
XX
KW Superantigen toxin; SAg; Staphylococcal enterotoxin B; SEB; cytostatic;
KW antibacterial; vaccine; MHC class II receptor; T-cell antigen receptor;
KW antibody; toxoid; staphylococcal/streptococcal toxin; diagnosis;
KW treatment; superantigen-associated bacterial infection; ds.
XX
OS Staphylococcus sp.
XX
FH Key   Location/Qualifiers
FT CDS   244. .1044
FT   /*tag= a
FT   /product= "Staphylococcal enterotoxin B"
FT sig_peptide   247. .327

```
FT                   /*tag=  b
FT   mat_peptide     328. .1041
FT                   /*tag=  c
FT                   /label= Mature_Staphylococcal_enterotoxin_ B
FT                   /note= "Also includes the ATG start codon"
FT                   /transl_except= (pos:772. .774, aa:Tyr)
FT                   /transl_except= (pos:901. .903, aa:Gln)
FT                   /transl_except= (pos:904. .909, aa:Asn)
FT                   /transl_except= (pos:1012. .1014, aa:Leu)
XX
PN   WO200009154-A1.
XX
PD   24-FEB-2000.
XX
PF   13-AUG-1998;   98WO-US016766.
XX
PR   13-AUG-1998;   98WO-US016766.
XX
PA   (REED-) REED ARMY INST RES WALTER.
XX
PI   Ulrich RG,  Olson MA,  Bavari S;
XX
DR   WPI; 2000-224177/19.
DR   P-PSDB; AAY70104.
XX
PT   Nucleic acid encoding superantigen toxin useful as a vaccine and for
PT   diagnosis of superantigen-associated bacterial infections.
XX
PS   Claim 4; Page 76-77; 118pp; English.
XX
CC   The present sequence is the DNA encoding Staphylococcal enterotoxin B
CC   (SEB), a bacterial superantigen toxin (SAg). The coding region of the SAg
CC   toxin when altered by site directed mutagenesis, results in disruption of
CC   binding of the toxin to both the MHC class II or T-cell antigen receptor.
CC   SEB has antibacterial and cytostatic activity. This sequence is useful
CC   for the production of SEB vaccines and specific antibodies. This vaccine
CC   overcomes the disadvantages of the chemically inactivated toxoids and is
CC   designed to protect individuals against one or several related
CC   staphylococcal and streptococcal toxins. It is used for the diagnosis and
CC   treatment or amelioration of superantigen- associated bacterial
CC   infections
XX
SQ   Sequence 1712 BP; 596 A; 223 C; 255 G; 638 T; 0 U; 0 Other;

  Query Match             10.0%;  Score 184.2;  DB 3;  Length 1712;
  Best Local Similarity   54.5%;  Pred. No. 8.1e-21;
  Matches  552;  Conservative    0;  Mismatches  408;  Indels   53;  Gaps    7;

Qy        621 ACCTATCTATTCGTAAAAAGATAAAAATAACTATTGTTTTTTTTGTTATTTTATAATAAA 680
              | |  | | ||  |||| |  |     ||  ||   ||   |    |||  |    |||
Db         29 AACACTATGTTGTTAAAGATGTTTTCGTATATAAGTTTAGGTGATGTATAGTTACTTAAT 88

Qy        681 ATTATTAATATAAGTTAATGTTTTTTAAAAATATACAATTTTATTCTATTTATAGTTAGC 740
              |||  |  |||| |||||    | | || || ||   |||  |    |||   ||  | |
Db         89 TTTAAAAGCATAACTTAATTAATATAAATAACATGAGATTATTAAATATAATTAAGTTTC 148

Qy        741 TATTTTTTCATTGTTAGTAATATTGGTGAATTGTAATAACCTTTTTAAATCTAGAGGAGA 800
              | ||  |   || ||| |   ||   | |  |   |  |  | ||||||        |||
Db        149 TTTTAATGTTTTTTTAATTGAATATTTAAGATTATAACATATATTTAAAGTGTATCTAGA 208

Qy        801 ACCCAGATATAAA--ATGGAGGAATATTAATGGAAAACAATAAAAAAGTATTGAAGAAAA 858
                |   |  ||  ||||  ||    |   |||   |||| | | |||| |     |
Db        209 TACTTTTTGGGAATGTTGGATAAAGGAGATAAAAAATGTATAAGAGATTATTTATTTCAC 268

Qy        859 TGGTATTTTTTGTTTT---AGTGACATTTCTTGGACTAACAATCTCGCAAGAGGTATTTG 915
                |||  ||||  | ||   | ||| |||  ||      |||  |         | |
Db        269 ATGTAATTTTGATATTCGCACTGATATTAGTTATTTCTACACCCAACGTTTTAGCAGAGA 328

Qy        916 CTCAACAAGACCCCGATCCAAGCCAACTTCACAGATCTAGTTTAGTTAA---AAACCTTC 972
              |||||  |||  ||  | |||     |  | |||| |||  |||  | | |       |
Db        329 GTCAACCAGATCCTAAACCAGATGAGTTGCACAAATCGAGTAAATTCACTGGTTTGATGG 388

Qy        973 AAAATATATATTTTCTTTATGAGGGTGACCCTGTTACTCACGAGAATGTGAAATCTGTTG 1032
                                                                          

                                                                          
Db        389 AAGATATGAAAGTTTTGTATGATGATAATCATGTATCAGCAATAAACGTTAAATCTATAG 448

Qy       1033 ATCAACTTTTATCTCACCATTTAATATATAATGTTTCAGGGCCAA---------ATTATG 1083
              |||||  ||  |||      | |||||||||  |  ||    |  | |       ||||||
Db        449 ATCAATTTCTATACTTTGACTTAATATATTCTATTAAGGACACTAAGTTAGGGGATTATG 508

Qy       1084 ATAAATTAAAAACTGAACTTAAGAACCAAGAGATGGCAACTTTATTTAAGGATAAAAACG 1143
              ||||  |   |  ||| |||| ||| ||||  | ||   |  ||  || ||||||  |||
Db        509 ATAATGTTCGAGTCGAATTTAAAAACAAAGATTTAGCTGATAAATACAAAGATAAATACG 568

Qy       1144 TTGATATTTATGGTGTAGAATATTACCATCTCTGTTATTTATGTGAAAA---------- 1192
              | ||| | | ||| |   | |||||  |||  |||||||| | | ||||
Db        569 TAGATGTGTTTGGAGCTAATTATTATTATCAATGTTATTTTTCTAAAAAAACGAATGATA 628

Qy       1193 ----------------TGCAGAAAGGAGTGCATGTATCTACGGAGGGGTAACAAATCATG 1236
                             ||  || | |  | ||||| || || || |||||  | |||
Db        629 TTAATTCGCATCAAACTGACAAACGAAAAACTTGTATGTATGGTGGTGTAACTGAGCATA 688

Qy       1237 AAGGGAATCATTTAGAAATTCCTAAAAAGATAGTCGTTAAAGTATCAATCGATGGTATCC 1296
              | || || || ||||| |   || ||  ||    |||  ||||      |||||||
Db        689 ATGGAAACCAATTAGATAAATATAGAAGTATTACTGTTCGGGTATTTGAAGATGGTAAAA 748

Qy       1297 AAAGCCTATCATTTGATATTGAAACAAATAAAAAAATGGTAACTGCTCAAGAATTAGACT 1356
              |    |||| ||||  |  |||| |||| |||| ||| ||||||||||||||||| |
Db        749 ATTTATTATCTTTTGACGTACAAACTAATAAGAAAAAGGTGACTGCTCAAGAATTAGATT 808

Qy       1357 ATAAAGTTAGAAAATATCTTACAGATAATAAGCAACTATATACTAATGGACCTTCTAAAT 1416
              |   | | | | ||| |    | |||||  |||| |||     |      ||    |
Db        809 ACCTAACTCGTCACTATTTGGTGAAAAATAAAAAACTCTATGAATTTAACAACTCGCCTT 868

Qy       1417 ATGAAACTGGATATATAAAGTTCATACCTAAGAATAAAGAAAGTTTTTGGTTTGATTTTT 1476
              |||||||  |||||||| || || |||    | ||| |  | || |||||||  ||| |
Db        869 ATGAAACGGGATATATTAAATTTATA---GAAAATGAGAATAGCTTTTGGTATGACATGA 925

Qy       1477 TCCCTGAACCAG------AATTTACTCAATCTAAATATCTTATGATATATAAAGATAATG 1530
              | |||| |||||      |||||  | |||||||||||  | |||||  || || || |||
Db        926 TGCCTGCACCAGGAGATAAATTTGCCCAATCTAAATATTTAATGATGTACAATGACAATA 985

Qy       1531 AAACGCTTGACTCAAACACAAGCCAAATTGAAGTCTACCTAACAACCAAGTAA 1583
              ||| | |||| || ||       | |||||||| || || || || ||| ||
Db        986 AAATGGTTGATTCTAAAGATGTGAAGATTGAAGTTTATCTTACGACAAAGAAA 1038

Search completed: May  9, 2004, 13:27:40
Job time : 706 secs
```